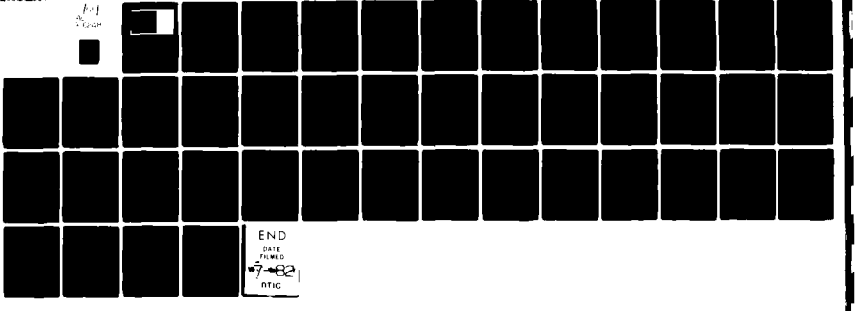
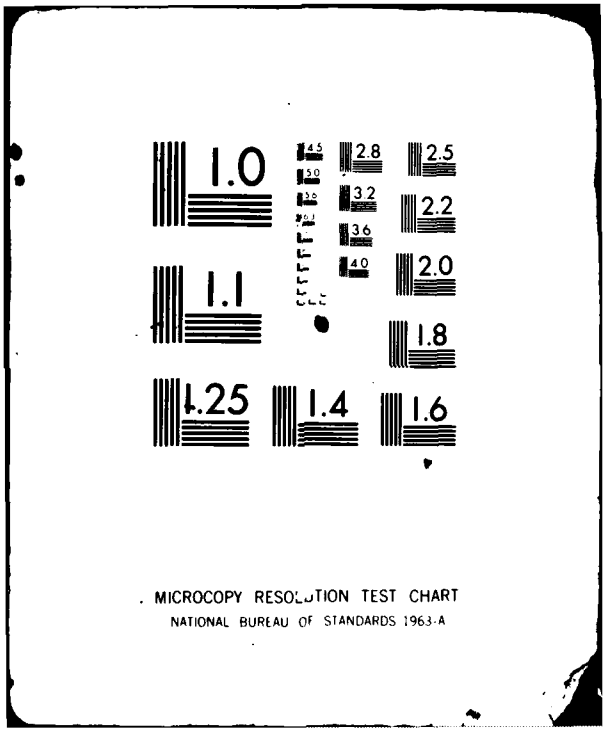
AD-A116 240 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
TOPICS IN MULTIVARIATE APPROXIMATION THEORY.(U)
MAY 82 C DE BOOR DAAG29-80-C-0041
UNCLASSIFIED MRC-TSR-2379 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

MRC Technical Summary Report #2379

TOPICS IN MULTIVARIATE APPROXIMATION THEORY

C. de Boor

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

May 1982

(Received April 27, 1982)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

TOPICS IN MULTIVARIATE APPROXIMATION THEORY

C. de Boor

Technical Summary Report #2379

May 1982

ABSTRACT

Lectures delivered at the S.R.C. Numerical Analysis Summer School and Workshop at the University of Lancaster, England, July 19 - August 20, 1981. Topics include tensor products, multivariate polynomial interpolation, esp. Kergin Interpolation, and the recent developments of multivariate B-splines.

AMS (MOS) Subject Classifications: 41-02, 41A05, 41A10, 41A15, 41A63, 41A65

Key Words: multivariate, B-splines, Kergin interpolation, linear projectors.

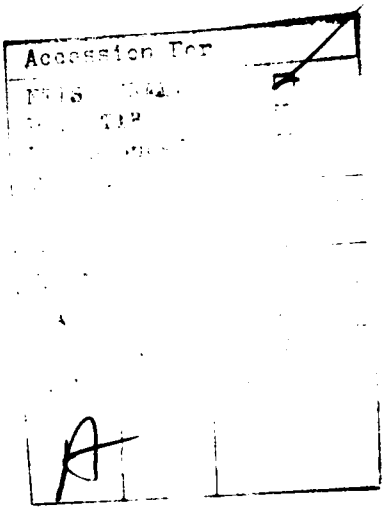
Work Unit Number 3 - Numerical Analysis and Computer Science

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

These lectures were intended to introduce young numerical analysts to some recent research topics in multivariate approximation theory. After lectures on approximation theory and linear approximation to set the stage, an introduction is given to the recent advances in multivariate B-splines and in multivariate polynomial interpolation.

These developments may well provide the theoretical foundation for efficient methods of multivariate smooth approximation in the future.



TOPICS IN MULTIVARIATE APPROXIMATION THEORY

C. de Boor

1. Approximation Theory

In this first lecture, I intend to give an overview of what is understood by the term "Approximation Theory". This is both a bow toward the title of these lectures and a survey of the kinds of things you might reasonably expect to see covered in these lectures, albeit with the special accent of "multivariate", but which I will for the most part not cover at all. In effect, this allows you to locate within the large scheme of things the few specific items I do cover.

Approximation Theory is usually understood to deal first and foremost with **best approximation**, or b.a. for short. This is the task of finding, given an element x of some **metric space** X, an element m^* from some given subset M of X for which

$$\operatorname{dist}(x,m^*) = \inf_{m\in M} \operatorname{dist}(x,m) =: \operatorname{dist}(x,M) .$$

Such an m^* is called a b.a. to x from M. In symbols:

$$m^* \in \mathcal{P}_M(x) .$$

Basic questions asked concern:

Existence: $|\mathcal{P}_M(x)| > 0$?

Uniqueness: $|\mathcal{P}_M(x)| < 2$? More generally, $|\mathcal{P}_M(x)| = ?$

Characterization: How would one recognize a b.a. (other than by the brute force approach of comparing it with all candidates)? This is particularly important for the next question.

Construction.

A priori bounds: What can be said about $\operatorname{dist}(x,M)$ based on the information that x lies in some set K ?

Details of the answers depend strongly on the specifics of X, dist, and M. Most commonly, X is a normed linear space, such as

$C(T)$:= continuous functions on some locally compact metric space T

(e.g., $T = [a,b]$ or $T = \mathbb{R}^n$ or whatever) and the metric is provided by the norm on X.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

Existence requires that M be closed. Beyond that, it is usually a matter of local compactness: A **minimizing sequence** (m_n) in M is picked; this means that $m_n \in M$ and

$$\lim_{n\to\infty} \|x - m_n\| = \operatorname{dist}(x,M) .$$

Then $\{m_n : n=1,2,\dots\}$ is bounded, hence, by local compactness, has a limit point m in M. For this (or any other) limit point,

$$\operatorname{dist}(x,M) \le \|x-m\| \le \limsup \|x-m_n\| = \lim \|x-m_n\| = \operatorname{dist}(x,M) ,$$

therefore $m \in P_M(x)$.

The standard example for M is a finite dimensional linear subspace of X, e.g.,

$$\pi_n := \text{polynomials of degree} \le n$$

as a subspace of $X = C[a,b]$ or $L_2[a,b]$. The desired local compactness is obvious for such an M. If M is a nonlinear subset, e.g., $M = \pi_n/\pi_m :=$ rational functions of degree n over m, the argument becomes more sophisticated: The convergence notion used is weakened sufficiently to gain local compactness while not losing the semicontinuity of the norm with respect to this notion of convergence.

A real difficulty in multivariate approximation is the fact that it becomes reasonable to consider infinite dimensional M. E.g.,

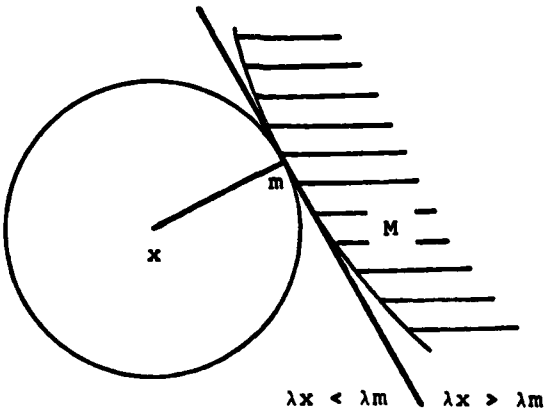
$$M = C[a,b] + C[c,d] \subseteq C([a,b]\times[c,d])$$

provides a simple example of the reasonable attempt to approximate a function of many variables (in this case, two) by composition of functions with fewer variables (in this case, the sum of two functions of one variable). Now even existence is a nontrivial matter.

Uniqueness and **characterization** involve a ball game, of sorts. Imagine the closed ball $\bar{B}_r(x)$ of radius r around x. Starting with $r = 0$, let r grow until $r = r^* := \operatorname{dist}(x,M)$. Then

$$P_M(x) = M \cap \bar{B}_{r^*}(x) .$$

For general M and some x, this first touch may well happen at two or more places. In such a circumstance, local uniqueness and characterization of a local(ly) b.a. become interesting questions. The interesting exception to this general statement is provided by a **convex** M, in which case we have the following picture:



Geometric fact. If M is convex and $m \in M$ and $r := \|x-m\|$, then $m \in P_M(x)$ iff it is possible to separate M and $B_r(x)$ by a closed hyperplane through m, i.e., iff there exists $\lambda \in X^*$ for which

$$\lambda[B_r(x)] < \lambda m \leq \lambda[M] \tag{1.1}$$

The first inequality in (1.1) is equivalent to

$$\lambda[B_r(0)] < \lambda(m-x) ,$$

i.e., to

$$\lambda \neq 0 \quad \text{and} \quad \|\lambda\| r \leq \lambda(m-x) \quad \left(\leq \|\lambda\| \|m-x\| = \|\lambda\| r \right) ,$$

hence to

$$\lambda \neq 0 \quad \text{and} \quad \lambda(m-x) = \|\lambda\| \|m-x\| .$$

This last condition is called variously ($\lambda \neq 0$ and) " $m-x$ is an extremal for λ " or " λ takes on its norm on $m-x$ " or " λ is parallel to $m-x$ ". I like this last phrase best and therefore write this condition

$$\lambda \parallel m-x$$

to remind you of the familiar picture: In a finite dimensional setting, λ would simply be the vector **normal** to the (separating) hyperplane and would point in the same direction as $m-x$ does; in short, λ would be parallel to $m-x$.

As a matter of convenience, one talks instead about the error $x-m$. This requires switching the sign of λ and so gives the

Characterization Theorem. Let M be a convex subset of the normed linear space X, let $x \in X \setminus \overline{M}$, and $m \in M$. Then

(i) $m \in \mathcal{P}_M(x) \iff \exists\, \lambda \parallel x-m$ s.t. $\lambda[M] \leq \lambda m$.

(ii) If M is a linear subspace, then

$m \in \mathcal{P}_M(x) \iff \exists\, \lambda \parallel x-m$ s.t. $\lambda \perp M$ (i.e., $\lambda[M] = \{0\}$).

Here is a good exercise which can be handled by entirely elementary means: Let $x \in X$, a nls, let $\lambda \in X^* \setminus 0$ and $m \in \ker \lambda$. Then, $m \in \mathcal{P}_{\ker \lambda}(x) \iff \lambda(x-m) = \|\lambda\|\|x-m\|$. For it, you might want to prove first that

$$x \in X,\ \lambda \in X^* \implies |\lambda x| = \|\lambda\| \operatorname{dist}(x, \ker \lambda) \tag{1.2}$$

which contains all the customary error estimates of elementary numerical analysis.

Since $\lambda \parallel x-m$ and $\lambda \perp M$ together imply that $\lambda \parallel x-m'$ for all $m' \in \mathcal{P}_M(x)$, nonuniqueness in case of a linear M is tied to the possibility of such a λ being parallel to more than one element (of the same size), i.e., for the hyperplane

$$\{y \in X : \lambda y = 1\}$$

to touch the unit ball $B_1(0)$ at more than one point. Since both M and $\overline{B}_{\operatorname{dist}(x,M)}(x)$ are convex, having two distinct points m and m' in $\mathcal{P}_M(x)$ implies that the whole line segment $[m,m']$ between m and m' is in $\mathcal{P}_M(x)$. This says that the unit ball must contain line segments in its boundary, which is the same as saying that the norm is not strictly convex. Put positively, strict convexity of the norm (such as the L_p-norm for $1 < p < \infty$) implies uniqueness of b.a. from a convex subset.

For more specific choices of X and M, the characterization theorem can be made more explicit, to the point where it can be used for the construction of a b.a. For example, if $X = C(T)$ and M is an n-dimensional linear subspace of X and $m \in M$, then

$m \in \mathcal{P}_M(x) \iff$

$$\exists\, r<n,\ (w_i),(t_i) \text{ s.t. } \lambda := \Sigma_1^{r+1} w_i[t_i] \parallel x-m \quad \text{and} \quad \lambda \perp M.$$

Here, $[t]$ denotes the linear functional of point evaluation at t, i.e., $[t]: f \longmapsto f(t)$. Behind this specialization of the general characterization theorem is the result useful for Numerical Analysis that any linear functional on an $(n+1)$-dimensional subspace of $C(T)$ has a norm preserving extension of the form $\Sigma_1^{n+1} w_i[t_i]$.

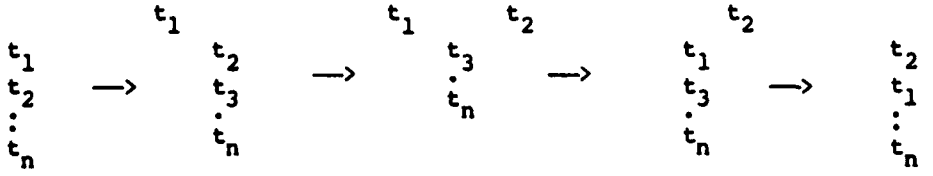
You will recognize in this characterization the familiar statement that the error $x-m$ in a b.a. must take on its norm at points t_1, ..., t_{r+1} with $r < n$ and such that, for some weights w_i with $w_i(x-m)(t_i) > 0$, all i, one has $\Sigma_1^{r+1} w_i[t_i] \perp M$.

It is not difficult to see that nonuniqueness is connected with having $r < n$ here. (Also, the norm is not strictly convex, so we would expect nonuniqueness for some x and M.) Recall that an n-dimensional subspace M of $C(T)$ is called a Haar space if, for any distinct points $t_1, \ldots, t_n$ in T, $([t_i])_1^n$ is linearly independent over M. For a Haar space M, having $0 \neq \lambda := \Sigma_1^{r+1} w_i[t_i] \perp M$ implies that $r \geq n$. Having in addition that $\lambda \| x-m$ implies that $\lambda \| x-m'$ for any other b.a. m', therefore m and m' must agree at the points $t_1, \ldots, t_{r+1}$ (assuming without loss that $w_i \neq 0$, all i), and, using once more that M is Haar, this implies that $m = m'$. Conversely, one can show that, if M is not Haar, then there are functions with many b.a.'s from M.

This equivalence between uniqueness and the Haar property has unhappy consequences for multivariate approximation, because of the following

Fact (Mairhuber). <u>If</u> T <u>is not essentially just an interval, and if</u> $\dim M > 1$, <u>then</u> M <u>is not Haar.</u>

The proof consists of a bit of railroading: Let $(f_1, \ldots, f_n)$ be a basis for M. Then $\det(f_j(t_i))$ is a continuous function of the n points $t_1, \ldots, t_n$. If now T contains a "Y", i.e., a "fork" or "switch", then one can continuously deform $(t_1, t_2, t_3, \ldots, t_n)$ into $(t_2, t_1, t_3, \ldots, t_n)$ while keeping the t_i's distinct:



This means that the determinant has changed sign along the way, hence must have vanished for some choice of n distinct points.

The resulting nonuniqueness of b.a. in $C(T)$ for multidimensional T has produced a great industry in uniform approximation by functions of several variables and much fun can be had. I shall resist the temptation to enter into details now, because I am not convinced that best approximation is all that practical.

The question of **a priori bounds** or **degree of approximation** is concerned with

$$\mathrm{dist}(K,M) := \sup_{x \in K} \mathrm{dist}(x,M)$$

A typical example would be : $X = C[0,1]$, $M = \pi_n$, and $K := \{ x \in X : \|x''\|_\infty \leq 1 \}$. Actually, it is not easy or even useful to be precise without getting simply the number $\mathrm{dist}(K,M)$. The question of degree of approximation comes into its own when one has given a **scale** (M_h) or (M_n) of subsets with $h \longrightarrow 0$ or $n \longrightarrow \infty$ and then considers

$$E_K(h) := \mathrm{dist}(K, M_h)$$

as a function of h . One proves **direct** or **Jackson**-type theorems:

$$x \in K \quad ===> \quad \mathrm{dist}(x, M_h) \sim h^r$$

and tries to demonstrate their sharpness, if possible, by proving **inverse** or **Bernstein**-type theorems:

$$\mathrm{dist}(x, M_h) \sim h^r \quad ===> \quad x/\|x\| \in K .$$

Related is the question: Given that $x \in K$, is the scale (M_h) a good choice for approximating x ? What is one to judge by? Kolmogorov [1936] proposed some time ago that $\mathrm{dist}(x,M)/\dim M$ is a good measure. He introduced

$$d_n(K) := \inf_{\dim M \leq n} \mathrm{dist}(K,M) =: \text{the } \textbf{n-width of } K .$$

While it is not easy to find an **optimal** subspace, i.e., a subspace at which the infimum of the definition is taken on, one would at least like an **asymptotically optimal** scale (M_n) , i.e., a scale with $\dim M_n = n$ for which $\mathrm{dist}(K,M_n) = O(d_n(K))$.

Once **effort** enters considerations of approximation (here in the form of the dimension of M , i.e., the degrees of freedom used in the approximation), one can, of course, ask more: Is it really worthwhile to construct **best** approximations, particularly when a near-best approximation is cheaply available? Here we call the linear map $A: X \longrightarrow M$ a **near-best** approximation scheme if

$$\exists \text{ const } \forall x \in X \quad \|x - Ax\| \leq \text{const } \mathrm{dist}(x,M) .$$

Here, const = 1 would be best possible. Ax would then be a b.a. for every x . This does happen in inner product spaces but hardly anywhere else except in very special circumstances. In any event, such A is necessarily a **linear projector** (with $\mathrm{ran}\, A = M$) since the inequality implies that $A_{|M} = 1$. These approximation schemes are the topic of the second lecture.

There are many books on Approximation Theory available. One of the most striking is Lorentz [1966] . Akhiezer [1967] summarizes the classical part. Both Cheney [1966] and Rivlin [1969] provide a careful modern introduction to the field while the two volumes of Rice [1964, 1969] bring quite a bit of additional material, especially on approximation from a nonlinear M . Powell [1981] and Schönhage [1971] each give a very interesting view of the subject.

2. Linear Interpolation

Linear projectors arise from interpolation, as I intend to make clear in this lecture. Fortunately for me, Ward Cheney is with us who has spent a good part of his professional life studying linear projectors. He will no doubt be ready to answer all questions left over after (or raised by) this lecture. Look for his publications (e.g., Morris & Cheney [1974]). See also the excellent book by Davis [1963].

The setup is quite simple: We have a linear space X of functions on some domain T and, correspondingly, the linear space X' of linear functionals on X. We have given $f_1, \ldots, f_m \in X$ and $\lambda_1, \ldots, \lambda_n \in X'$ and consider the

Task: Given $g \in X$, construct $Pg := \Sigma_1^m \alpha(j) f_j$ such that Pg **interpolates to** g **at** $\lambda_1, \ldots, \lambda_n$, i.e.,

$$\lambda_i Pg = \lambda_i g, \quad i=1,\ldots,n.$$

For example, the specifics: $T = [a,b]$, $X = C[a,b]$, $f_j = (\,)^{j-1}$ and $\lambda_i = [t_i]$ describe the task of polynomial interpolation. Altering this to $\lambda_i : f \longmapsto \int_T f(x) f_i(x) dx$, all i, and choosing $m = n$ describes least-squares approximation by polynomials.

Our first observation is that this task does not depend on the individual functions $f_1, \ldots, f_m$ nor on the individual linear functionals $\lambda_1, \ldots, \lambda_n$, but only on their spans

$$F := \operatorname{span}(f_i)_1^m := \{\Sigma_1^m \alpha(j) f_j : \alpha \in \mathbb{R}^m\}$$

and

$$\Lambda := \operatorname{span}(\lambda_i)_1^n .$$

This is obvious for the f_i's since the very task is stated in terms of their span. As to Λ, observe that

$$\lambda_i g = \lambda_i h, \ i=1,\ldots,n \iff \forall \beta \in \mathbb{R}^n \ (\Sigma \beta(i)\lambda_i) g = (\Sigma \beta(i)\lambda_i) h .$$

We use the abbreviation

$$\mathrm{LIP}(F,\Lambda)$$

for the **Linear Interpolation Problem given by** F **and** Λ, i.e., for the

Task: Given $g \in X$, find $Pg \in F$ s.t. $g = Pg$ on Λ.

Here, F and Λ are understood to be linear subspaces (finite dimensional) of X and X', respectively. We call $\mathrm{LIP}(F,\Lambda)$ **correct** if the task has exactly one solution for every $g \in X$.

Now, having just gotten rid of the f_i's and λ_i's, it is convenient to reintroduce them, in a possibly refined form: Let $(f_i)_1^m$ be a **basis** for F and let $(\lambda_i)_1^n$ be a **basis** for Λ. Then

$\Sigma_j \alpha(j) f_j$ solves LIP(F,Λ) for given g <===>

$\alpha \in R^m$ solves $\Sigma_j (\lambda_i f_j)\alpha(j) = \lambda_i g$, $i=1,\ldots,n$.

We conclude

Lemma. (i) LIP(F,Λ) <u>is correct</u> <===> <u>the</u> **Gramian** $G := (\lambda_i f_j)_{i=1,j=1}^{m\ \ n}$ <u>is invertible</u>.

(ii) LIP(F,Λ) <u>is correct</u> ===> $Pg = \Sigma_j \alpha(j) f_j$ <u>with</u> $\alpha = A^{-1}(\lambda_i g)$.

The **proof** is linear algebra: Since (f_i) is linearly independent, uniqueness is equivalent to having A 1-1. Since (λ_i) is linearly independent, existence is equivalent to having A onto. Note that correctness implies $m = n$.

The map P defined by such a correct LIP is **linear** (as a composition of the linear maps $g \longmapsto (\lambda_i g) \longmapsto \alpha = G^{-1}(\lambda_i g) \longmapsto \Sigma_j \alpha(j) f_j$). Also, by uniqueness, $P_{|F} = 1$, hence $P^2 = P$, showing that P is a **linear projector.** Its range is

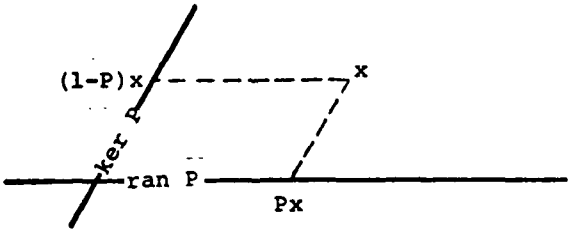
$$\text{ran } P = F = \{x \in X : Px = x\}$$

while its kernel or nullspace is

$$\ker P = \{x \in X : \lambda x = 0, \text{ all } \lambda \in \Lambda\} =: \Lambda_\perp = \text{ran}(1-P).$$

The customary view of a linear projector is that it provides a **direct sum decomposition:**

$$x = Px + (1-P)x .$$



Px is the projection of x onto ran P along ker P. I prefer to think of P as **given by** $F = \text{ran } P$ **and its interpolation conditions**

$$\Lambda = \{\mu \in X' : \mu P = \mu\} = \text{ran } P' .$$

This stresses the fact that Px is the unique element in ran P which agrees with x on Λ.

The **construction** of Pg involves, off hand, the solution of the linear system $G\alpha = (\lambda_i g)$. This can be viewed as switching over to the

new basis (λ_i') for Λ with

$$\lambda_i' := \Sigma_j G^{-1}(i,j)\lambda_j \ .$$

Such a basis is, by its construction, dual to (f_i), i.e.,

$$\lambda_i' f_j = \delta_{ij} \ .$$

Another possibility is the **Lagrange** approach: Switch over to

$$f_j' := \Sigma_i G^{-1}(i,j) f_i$$

so that now $(\lambda_i f_j') = 1$. Yet another possibility is the **Newton** approach: If possible, switch over to

$$f_j' := \Sigma_i R^{-1}(i,j) f_i \quad , \quad \lambda_i' := \Sigma_j L^{-1}(i,j)\lambda_j$$

with LR a **triangular factorization** for G, giving again $(\lambda_i' f_j') = 1$. In this last approach, we would need $(\lambda_i f_j)_1^r$ invertible for $r=1,2,\ldots$. Equivalently, we would need the $LIP(F_r,\Lambda_r)$ with $F_r := \text{span}(f_1,\ldots,f_r)$, $\Lambda_r := \text{span}(\lambda_1,\ldots,\lambda_r)$ to be correct, giving rise to a projector P_r, $r=1,2,\ldots,n$. In these terms, suitable bases for F and Λ can be constructed bootstrap fashion:

$$f_i' = (1-P_{i-1})f_i \ , \quad \lambda_i' = \lambda_i(1-P_{i-1})$$

$$P_i = \sum_{j<i} f_i'\lambda_i' \,/\lambda_i' f_i'$$

which is, in effect, Gauss elimination without pivoting.

The **example** inspiring all this terminology is, of course, **polynomial interpolation**, mentioned earlier, in which $f_j = (\,)^{j-1}$ and $\lambda_i = [t_i]$. The Lagrange approach leads to

$$f_j' : t \longmapsto \Pi_{i\neq j}(t-t_j)/(t_i-t_j)$$

while the Newton approach leads to

$$f_j' : t \longmapsto \Pi_{i<j}(t-t_j) \ , \quad \lambda_i' = [t_1,\ldots,t_i] :=$$

the **divided difference** at $t_1, \ldots, t_i$, and thence to the Newton form

$$Pg = \Sigma_1^n [t_1,\ldots,t_i]g \ \prod_{j<i}(\cdot - t_j) \ .$$

Another well known example is specified by : $X = C(T)$, $F = \text{span}(f_i)$, of dimension n, and

$$\lambda_i : g \longmapsto \int_T f_i g \ , \ i=1,\ldots,n \ .$$

Now P is **Least-squares** approximation, and the Newton approach is, in this instance, called **Gram-Schmidt orthogonalization.**

We are interested in linear projectors because they provide near-best linear approximation schemes. Explicitly, we have

Lebesgue's Inequality: $\|g - Pg\| \leq \|1-P\| \ \text{dist}(g,F)$

in case X is a normed linear space (=: nls) and P is bounded. In

fact, we have a bit more:

$$|\mu g - \mu P g| \le \operatorname{dist}(\mu, \Lambda)\, \|1-P\|\, \operatorname{dist}(g, F)\ , \quad \mu \in X^*,\ g \in X\ . \tag{2.1}$$

For the proof, note that, for $\lambda \in \Lambda$ and $f \in F$, $\lambda(1-P) = 0$, $(1-P)f = 0$, so

$$|\mu g - \mu P g| = |(\mu-\lambda)(1-P)(g-f)| \le \|\mu-\lambda\|\,\|1-P\|\,\|g-f\|\ .$$

Now take the infimum over $\lambda \in \Lambda$ and $f \in F$.

The inequality (2.1) is important for the rule makers who customarily approximate μg by $\mu P g$ ever since Newton proposed this for $\mu g = \int g$ and P polynomial interpolation. The variational approach to splines, particularly important for the understanding of Duchon's multivariate 'thin plate' splines (Duchon [1976], [1977], Meinguet [1979]), takes off from this setup. See the epilogue.

The basic inequality (2.1) raises the two questions:
(i) Is P bounded? (ii) How big is $\|P\|$ (or, $\|1-P\|$)?
The following two lemmas give answers of sorts.

Lemma 2.1. Let P be given on the nls X by LIP(F,Λ). Then P is bounded iff $\Lambda \subseteq X^*$ (:= continuous linear functionals on X).

Proof. "==>" For all $\lambda \in \Lambda$, $\lambda = \lambda P = (\lambda_{|F}) \circ P$ with $\lambda_{|F}$ continuous since $\dim F < \infty$.

"<==" We can write $P = \Sigma_1^n f_i \lambda_i$, hence $\|P\| \le \Sigma_1^n \|f_i\|\,\|\lambda_i\| < \infty$. |||

Lemma 2.2. $\|P\| = \sup_{f \in F} \inf_{\lambda \in \Lambda} \|\lambda\|\,\|f\| / |\lambda f|$.

Proof. For any linear map A,

$$\|A\| = \sup_{x \notin \ker A} \|Ax\|/\|x\| = \sup_{x \notin \ker A} \sup_{y \in \ker A} \|Ax\|/\|x-y\|$$

$$= \sup_{x \notin \ker A} \|Ax\|/\operatorname{dist}(x, \ker A) = \sup_{x \notin \ker A} \inf_{\lambda \perp \ker A} \|Ax\|\,\|\lambda\|/|\lambda x|\ .$$

If now A is the linear projector P, then $x = Px + (1-P)x$ and $Px \in F = \operatorname{ran} P$ while $(1-P)x \in \ker P$. Further, $\lambda \perp \ker P$ iff $\lambda \in \Lambda$ (= ran P') , using the fact that $\operatorname{ran} P = \{x : x = Px\}$, hence $\lambda \perp \ker P$ ==> $\lambda \perp \operatorname{ran}(1-P)$ ==> $\lambda = \lambda P$ ==> $\lambda \in \operatorname{ran} P' = \Lambda$. Therefore,

$$\|P\| = \sup_{x \notin \ker P} \inf_{\lambda \perp \ker P} \|Px\|\,\|\lambda\|/|\lambda x| = \sup_{x \in F} \inf_{\lambda \in \Lambda} \|x\|\,\|\lambda\|/|\lambda x|\ . \ |||$$

3. The tensor product construct

Our first foray into a multivariate setup is by tensor products. This construct is of limited use. Yet when it can be employed, it is so efficient that it is worth some effort to bring a given approximation problem into this form, if it can be done at all. Somewhat surprising approximation theoretic advantages of tensor products are discussed in de Boor & DeVore [1981].

The mathematics is quite simple, yet papers still appear which look forbidding and needlessly complicated since they do not make full use of the fact that, when dealing with tensor products, everything is essentially univariate, even the computer programs.

Naively, the tensor product of two univariate linear approximation schemes or projectors P and Q is formed as follows. For each fixed y, the linear projector P is applied to the **y-section**

$$h_y := h(\cdot,y)$$

of the bivariate function h, giving the univariate function

$$Ph_y = \Sigma_j \alpha(j;h_y)f_j$$

in which the coefficients $\alpha(j) = \alpha(j;h_y)$ depend on h_y, hence on y. Then Q is applied to each of these coefficient functions

$$c_j : y \longmapsto \alpha(j;h_y) ,$$

thus obtaining their univariate approximations

$$Qc_j = \Sigma_k \beta(k;j)g_k ,$$

with (g_k) a basis for ran Q. Altogether, this gives the approximation

$$(Rh)(x,y) = \sum_{j,k} \beta(k;j)g_k(y)f_j(x)$$

to h.

Several questions are immediate: What is the nature of R? Is R a projector? What is the corresponding LIP? If we first use Q on each x-section of H and then P on the resulting coefficient functions, would the resulting approximation be again Rh?

We consider the last question first, since its answer supplies also the answer to the other questions. At its root is the question of whether the two operators $P\otimes 1$ and $1\otimes Q$ defined by

$$(P\otimes 1)h : (x,y) \longmapsto (Ph_y)(x) , \quad (1\otimes Q)h : (x,y) \longmapsto (Qh_x)(y)$$

commute. Since both P and Q are linear projectors, this question is settled once we know that $\lambda\otimes 1$ and $1\otimes\mu$ commute for each $\lambda \in \text{ran } P'$ and each $\mu \in \text{ran } Q'$. This is dealt with next in careful (and perhaps boring) detail.

Suppose that x_r is a ls of functions on some domain T_r, r=1,2. The tensor product $X_1 \otimes X_2$ of the two linear spaces X_1 and X_2 is customarily defined as the dual of the linear space of all bilinear functionals on $X_1 \times X_2$. Since I intend to use this concept only in the context of function spaces, I prefer the following definition: For $x_r \in X_r$,

$$x_1 \otimes x_2 : T_1 \times T_2 \longrightarrow \mathbb{R} : (t_1,t_2) \longmapsto x_1(t_1)x_2(t_2)$$

defines a function on $T_1 \times T_2$ called the **tensor product** of x_1 and x_2. Further,

$$X_1 \otimes X_2 := \text{span}\{ x_1 \otimes x_2 : x_r \in X_r\} \subseteq \mathbb{R}^{T_1 \times T_2} .$$

It is easy but essential to verify that the map

$$X_1 \times X_2 \longrightarrow X_1 \otimes X_2 : (x_1,x_2) \longmapsto x_1 \otimes x_2$$

is bilinear. This implies that

$$X_1 \otimes X_2 = \{ \Sigma_j x_{1j} \otimes x_{2j} : x_{rj} \in X_r\}$$

(i.e., the scalars in the linear combinations making up $X_1 \otimes X_2$ can be absorbed).

Prime examples are the spaces

$P_m \otimes P_n$:= polynomials in two variables, of degree < m in the first and of degree < n in the second,

and

$$\mathbb{R}^m \otimes \mathbb{R}^n = \mathbb{R}^{m \times n} = \text{m-by-n matrices} .$$

Correspondingly, one defines the tensor product $\lambda \otimes \mu$ of $\lambda \in X_1'$ and $\mu \in X_2'$ as a linear functional on $X_1 \otimes X_2$ by the rule

$$\lambda \otimes \mu : X_1 \otimes X_2 \longrightarrow \mathbb{R} : \Sigma_j u_j \otimes v_j \longmapsto \Sigma_j (\lambda u_j)(\mu v_j) .$$

This is so obviously a linear functional on $X_1 \otimes X_2$ that it is easy to overlook the only nontrivial (though certainly not very deep) point about this definition, viz. whether it is a definition. The problem is that the rule for the value of $\lambda \otimes \mu$ at $w = \Sigma_j u_j \otimes v_j$ makes explicit use of the particular representation of w mentioned. Elements of $X_1 \otimes X_2$ have many different representations. For example, if $u = a+b$ and $v = 2c$, then

$$u \otimes v = a \otimes v + b \otimes v = a \otimes (2b) \otimes c = a \otimes c + a \otimes c + b \otimes v = \ldots$$

We should therefore have, correspondingly, that

$$\begin{aligned} (\lambda u)(\mu v) &= (\lambda a)(\mu v) + (\lambda b)(\mu v) = \\ &= (\lambda a)(\mu v) + (2\lambda b)(\mu c) = \\ &= (\lambda a)(\mu c) + (\lambda a)(\mu c) + (\lambda b)(\mu v) = \ldots \end{aligned}$$

We must therefore verify that <u>the number</u> $\Sigma_j (\lambda u_j)(\mu v_j)$ <u>depends only on</u> λ, μ, <u>and</u> $w := \Sigma_j u_j \otimes v_j$, <u>and not on the particular representation for</u> w <u>used</u>, i.e., not on the individual u_j's and v_j's. For this,

let

$$v := \Sigma_j(\lambda u_j)v_j \ .$$

Then $v \in X_2$, and

$$v(t_2) = (\Sigma(\lambda u_j)v_j)(t_2) = \Sigma(\lambda u_j)v_j(t_2)$$

$$= \lambda(\Sigma u_j v_j(t_2)) = \lambda w(\cdot,t_2)$$

showing that v is the λ-section of w ,

$$v = w_\lambda \ .$$

In particular, v depends only on w and λ . On the other hand, $\Sigma_j(\lambda u_j)(\mu v_j) = \mu w_\lambda$.

Note that we get by symmetry that also $\Sigma_j(\lambda u_j)(\mu v_j) = \lambda w_\mu$, showing the hoped-for commutativity.

We are ready to define the tensor product of two LIPs (did the Tiny Tim craze ever catch on here?): Assume that $F_r \subseteq X_r \subseteq \mathbb{R}^{T_r}$ and $\Lambda_r \subseteq X_r'$ give rise to the correct $LIP(F_r,\Lambda_r)$ with corresponding invertible Gramian

$$G_r := (\lambda_{ri}f_{rj})_{i,j=1}^{n_r} \ .$$

Pick some space W of functions on $T_1 \times T_2$ which contains $F := F_1 \otimes F_2$. Further, pick some $\nu_{ij} \in W'$ for which

$$\nu_{ij|F} = (\lambda_{1i}\otimes\lambda_{2j})_{|F} \ , \text{ all } i,j \ ,$$

and set $\Lambda := \text{span}(\nu_{ij})$.

Then: (i) $(\bar{f}_{ij}) := (f_{1i}\otimes f_{2j})$ is a basis for F .

(ii) The $LIP(F,\Lambda)$ is correct.

(iii) The corresponding interpolant Pw to $w \in W$ can be computed as

$$Pw = \sum_{i,j} \Gamma_w(i,j)\,\bar{f}_{ij}$$

with

$$\Gamma_w = G_1^{-1}L_w(G_2^{-1})^T \quad \text{and} \quad L_w(i,j) := \nu_{ij}w \ , \text{ all } i,j \ . \tag{3.1}$$

For the proof, any $w \in F$ can be written $w = \Sigma_{ij}\Gamma(i,j)\bar{f}_{ij}$ for some suitable coefficient matrix Γ . From this, we compute that

$$L_w(r,s) := \nu_{rs}w = \sum_{i,j}\Gamma(i,j)(\lambda_{1r}f_{1i})(\lambda_{2s}f_{2j})$$

$$= \sum_{i,j} G_1(r,i)\,\Gamma(i,j)\,G_2(s,j) = G_1\,\Gamma\,G_2^T \ .$$

This shows that $\Gamma = G_1^{-1}L_w(G_2^T)^{-1}$, i.e., Γ depends only on w . This proves (i). Further, it shows that $(\lambda_{ij}w) = L$ <===> $\Gamma = G_1^{-1}L(G_2^T)^{-1}$, thus proving (ii) and (iii).

A simple example is provided by the tensor product of cubic Hermite interpolation, i.e., cubic interpolation at the four points $a, a, b, b,$ with osculatory parabolic interpolation, i.e., parabolic interpolation at the three points c, c, d. The above description leaves considerable freedom in the choice of the ν_{ij}. A natural choice would be

$$\nu_{ij}w := \begin{array}{c|ccc} i\backslash j & 1 & 2 & 3 \\ 1 & w(a,c) & w_y(a,c) & w(a,d) \\ 2 & w_x(a,c) & w_{xy}(a,c) & w_x(a,d) \\ 3 & w(b,c) & w_y(b,c) & w(b,d) \\ 4 & w_x(b,c) & w_{xy}(b,c) & w_x(b,d) \end{array}$$

and this would require the function space W to consist of functions defined (at least) on the rectangle $[a,b]\times[c,d]$ (assuming that $a < b$ and $c < d$) and to have first derivatives and the first cross derivative (in a pointwise sense). Further, the desired commutativity would require that $w_{xy} = w_{yx}$ for all $w \in W$.

The computational advantage of this construct over other means of approximation is considerable. It generally takes $O(N^3/3)$ operations to solve the linear system for the coefficients of the solution of a LIP using N degrees of freedom. This number can be reduced considerably in a univariate setup (such as in spline approximation) through the use of special bases which make the resulting system banded. This is much harder in a multivariate context. In any event, if the $\mathrm{LIP}(F_r,\Lambda_r)$ involves n_r degrees of freedom, $r=1,2$, then their tensor product involves n_1n_2 degrees of freedom. Yet, using (3.1), one can compute the requisite n_1n_2 coefficients in $O(n_1^3/3 + n_1^2n_2 + n_1n_2^2 + n_2^3/3)$ operations, since it only involves solving n_2 systems of order n_1 with the same coefficient matrix, and n_1 systems of order n_2 with the same coefficient matrix. Further savings are possible because this reduction of the computations to the univariate context provides the opportunity to make use of whatever savings are available in that context. If, for example, the univariate schemes give rise to banded systems, then their tensor product can be constructed in $O(n_1n_2)$ operations.

This leads to a point made in de Boor [1979], that it is possible to form the "tensor product" of the computer programs for the solution of the univariate "factor"-problems. Typically, one has available programs $\mathrm{INTER}_r(B,M,A)$ which take the input vector $B = (\lambda_{ri}g)_1^{n_r}$ with $M = n_r$, and produce from it the desired coefficient vector $A = \alpha$ for the interpolant $P_rg = \Sigma\, \alpha(j)f_{rj}$ to g. Two changes are required to

make such a program amenable to tensor product computations: One extends it to allow B to be a collection of N input vectors, i.e., to allow B to be an M-by-N matrix having these N data vectors as its columns; this requires that also N be supplied on input. Correspondingly, the output A becomes a collection of N coefficient vectors, i.e., a matrix. But, for reasons that will be obvious in a moment, the resulting coefficient vectors should form the rows of that matrix, i.e., A should be an N-by-M matrix. This gives the extension $\text{INTER}_r'(B,M,N,A)$. With it, we can solve the tensor product of the two $\text{LIP}(F_r,\Lambda_r)$, r=1,2 , by the two calls

$\text{CALL INTER}_1'(L_w,n_1,n_2,A)$

$\text{CALL INTER}_2'(A,n_2,n_1,\Gamma)$

The two programs could even be the same as, e.g., in bicubic spline interpolation. Further, once such extended programs have been written, it is possible to carry out tensor product interpolation using more than two factors. Finally, this formation of program "tensor products" is also helpful in the evaluation or other manipulation of the interpolant. In any event, the detailed programming effort takes place at the univariate level, just as in the mathematical analysis of the construct.

The error is easily obtained formally. Writing $E_r := 1 - P_r$ for the error in the linear approximation scheme P_r , we find

$$1 = (P_1+E_1)\otimes(P_2+E_2) = P_1\otimes P_2 + P_1\otimes E_2 + E_1\otimes P_2 + E_1\otimes E_2 . \quad (3.2)$$

This shows the error operator for the tensor product scheme $P_1\otimes P_2$ to be a sum of the univariate errors. The order of approximation is therefore no better than the worse of the two univariate schemes.

Now note that $P_rE_r = E_rP_r = 0$. This implies that any sum of terms from the right hand side of (3.2) gives a linear projector. In particular, Gordon [1969]$_{1,2}$ has proposed the use of the socalled **Boolean sum**

$$P_1\oplus P_2 := P_1\otimes 1 + 1\otimes P_2 - P_1\otimes P_2 = P_1\otimes P_2 + P_1\otimes E_2 + E_1\otimes P_2 .$$

The resulting approximation scheme is called **blending** since it uses interpolation conditions of the form $\lambda\otimes 1$ and $1\otimes\mu$, hence, in its simplest form, matches information along certain lines parallel to the axes and so constructs a surface by "blending" together certain curves. For blending, the error is the **product** of the univariate errors. This improvement over the tensor product is bought at a high price: An infinite amount of information about the function to be approximated is required. Gordon has dealt successfully with this problem by proposing

that one first use a relatively dense but finite amount of information to construct good approximations to the required curves and then use these approximations in the final construct.

4. Multivariate polynomial interpolation

I begin with a review of the standard notation concerning polynomials in m variables. The notation is designed to make it all look just as in the univariate case. The general **polynomial of total degree** $< k$ is , by definition, any linear combination

$$x \longmapsto \Sigma_{|\alpha|<k} A_\alpha x^\alpha$$

of the **monomials** $()^\alpha$ with $|\alpha| < k$. Here,

$$x^\alpha := x(1)^{\alpha(1)} \cdot \ldots \cdot x(m)^{\alpha(m)}$$

and the **length** $|\alpha|$ of the integer vector α is defined by

$$|\alpha| := \|\alpha\|_1 = \alpha(1) + \ldots + \alpha(m)$$

if, as we assume, all the components of α are nonnegative. For such an **index vector**, one sets

$$\alpha! := \alpha(1)! \cdot \ldots \cdot \alpha(m)!$$

and thereby recovers the binomial formula

$$(x + y)^\alpha = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} x^\alpha y^\beta .$$

The partial **ordering** used here is componentwise:

$$\beta \le \alpha := \text{for all } i,\ \beta(i) \le \alpha(i) .$$

This gives **Leibniz' formula**

$$D^\alpha(fg) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} (D^\beta f)(D^{\alpha-\beta} g)$$

for the derivative of a product. Here,

$$D^\alpha := D_1^{\alpha(1)} \ldots D_m^{\alpha(m)}$$

with D_i the partial derivative with respect to the ith argument. More generally, $p(D)$ is the constant coefficient differential operator

$$p(D) := \Sigma_\alpha A_\alpha D^\alpha$$

in case p is the polynomial

$$p = \Sigma_\alpha A_\alpha ()^\alpha .$$

For the special linear polynomial

$$p: x \longmapsto x^* y := \Sigma_i x(i) y(i) ,$$

we write

$$D_y := \Sigma\, y(i) D_i$$

instead of D^*y for the resulting (unnormalized) derivative in the direction of y.

In one variable, it is convenient to talk about P_k, the linear space of polynomials of order k, i.e., of degree $< k$, since its dimension is k and the optimal approximation order from P_k achievable on an interval of length h is h^k. In several variables, approximation order continues to be linked to (total) polynomial order, but the dimension and other interesting quantities are more easily expressed in terms of (total) degree rather than order. For this reason, I will concentrate on the linear space

$$\pi_k = \pi_k(\mathbb{R}^m)$$

of all polynomials of (total) degree $< k$ in m variables. It is not difficult to see that

$$\dim \pi_k(\mathbb{R}^m) = \binom{m+k}{m} .$$

Indeed, the rule

$$\alpha(r) := i(r) - i(r-1) - 1 , \quad r=1,\dots,m$$

sets up a 1-1 correspondence between

$$\{\alpha \in \mathbb{Z}_+^m : |\alpha| < k\}$$

and the set

$$\{I \subseteq \{1,\dots,m+k\} : |I| = m\}$$

of cardinality $\binom{m+k}{m}$ if we let $i(1), \dots, i(m)$ be the elements of I, in increasing order (and take $i(0) = 0$). Thus the generating 'sequence'

$$(()^\alpha)_{|\alpha|<k}$$

for π_k contains $\binom{m+k}{m}$ terms. On the other hand, this sequence is linearly independent since, e.g.,

$$[0]D^\beta()^\alpha = \alpha!\, \delta_{\beta\alpha} .$$

Note that

$$\dim (\pi_{k+1} \ominus \pi_k) = \binom{m-1+k}{m-1}$$

since $\{\alpha \in \mathbb{Z}_+^m : |\alpha| = k\}$ is in obvious 1-1 correspondence with $\{\alpha \in \mathbb{Z}_+^{m-1} : |\alpha| < k\}$. This reaffirms the well known identity

$$\binom{m+k}{m} = \sum_{r=0}^{k} \binom{m-1+r}{m-1} .$$

We now consider the $\mathrm{LIP}(\pi_k,T) := \mathrm{LIP}(\pi_k,\mathrm{span}([t])_{t\in T})$ with T a subset of $\mathbb{R}^m$. We call T correct (for π_k) if the $\mathrm{LIP}(\pi_k,T)$ is correct. Since π_k is not the tensor product of univariate polynomial spaces, it seems unlikely that we could employ the tensor product construct to obtain correct T's. Yet it is possible, as the following example, two-dimensional for simplicity, shows. Recall that the linear

projector of polynomial interpolation at points $u_0, \ldots, u_k$ can be written in Newton form as

$$P_u = \Sigma_0^k \phi_{ui}[u_0,\ldots,u_i]$$

with

$$\phi_{ui}(x) := (x-u_0)\ldots(x-u_{i-1}) \text{ , all } i .$$

Therefore, any partial sum

$$R_I := \sum_{(i,j)\in I} \phi_{ui}\otimes\phi_{vj}[u_0,\ldots,u_i]\otimes[v_0,\ldots,v_j]$$

of its tensor product with P_v for some point sequence $v_0, \ldots, v_k$ is also a linear projector, given that $I \subseteq \{0,\ldots,k\}^2$. The range of R_I is somewhere in $\pi_k\otimes\pi_k$. To insure that it is actually $\pi_k(\mathbb{R}^2)$, choose $I = \{(i,j) : i+j < k\}$. With this choice, ran $R_I \subseteq \pi_k$ and equality must hold since, by just counting terms, we see that ran R_I has dimension $\binom{2+k}{k}$ which is dim $\pi_k(\mathbb{R}^2)$. It is now a nice exercise to verify that, for this choice of I , R_I solves the LIP(π_k,T) with $T := \{(u_i,v_j) \in \mathbb{R}^2 : i+j<k\}$.

The same construction works in m variables and so provides the only standard choice of correct point sets T for $\pi_k(\mathbb{R}^m)$. This is the simplicial choice which, up to a linear change of variables, is

$$T = \{(u_p,v_q,\ldots,w_r) \in \mathbb{R}^m : p+q+\ldots+r<k\}$$

with (u_p), (v_p), ..., (w_p) given sequences of real numbers. Note that an affine change of variables

$$x \longmapsto Ax + b$$

leaves π_k invariant, hence leaves invariant the collection of correct point sets T for π_k .

More general correct point sets can be generated with the aid of the following theorem due to Chung & Yao [1977], - and here I must thank Dr. A. Genz for pointing out this reference to me.

Theorem 4.1. *If the point set $T \subset \mathbb{R}^m$ has cardinality dim $\pi_k(\mathbb{R}^m)$, and, for every $t \in T$, there exist k distinct hyperplanes on which all points in T lie except for t , then T is correct for $\pi_k(\mathbb{R}^m)$.*

It is clear how one would prove this theorem: For each $t \in T$, we can find, by assumption, k m-vectors a_1 , ..., a_k and scalars b_1 , ..., b_k so that the k-th degree polynomial

$$L_t : x \longmapsto (a_1{}^*x-b_1)\ldots(a_k{}^*x-b_k)$$

vanishes on $T\setminus t$ but not at t . This implies that, for any given g , the function

$$\Sigma_{t\in T}\, g(t)\, L_t/L_t(t)$$

is a polynomial of degree $< k$ which agrees with g on T. On the other hand, the 'sequence' $(L_t)_{t\in T}$ is linearly independent (since it is obviously independent over T), and, by assumption, contains exactly $|T| = \dim \pi_k$ terms, hence must be a basis for π_k and this establishes the uniqueness of the interpolating polynomial. In short, we have the generalization of Lagrange's way of treating univariate polynomial interpolation.

A particularly striking instance of such a correct set T are the "natural lattices" of Chung & Yao [1977] rediscovered recently by Dahmen & Micchelli [1980], and also by Hakopian [1981]$_2$. Pick n points $a_1, \ldots, a_n$ in $\mathbb{R}^m$ so that the points $0, a_1, \ldots, a_n$ are **in general position**, i.e., any $m+1$ of them are affinely independent. To recall, $m+1$ points $b_0, \ldots, b_m$ in $\mathbb{R}^m$ are **affinely independent** if their affine hull is all of $\mathbb{R}^m$, i.e., if $\mathrm{vol}_m \operatorname{conv} (b_i)_0^m \neq 0$. Then, for any subset I of $\{1,\ldots,n\}$ with $|I| = m$, there exists exactly one x_I for which

$$1 + a_i * x_I = 0 \text{ , all } i \in I .$$

(Indeed, since $0, (a_i)_{i\in I}$ are affinely independent, the sequence $(a_i)_{i\in I}$ must be linearly independent.) Further, for this x_I, we must have

$$1 + a_j * x_I \neq 0 \text{ , all } j \notin I$$

(since $1 + a_j * x_I = 0$ implies that $(a_i)_{i\in I}, a_j$ all lie in the hyperplane $\{x \in \mathbb{R}^m : 1 + x \cdot x_I = 0\}$, hence are not affinely independent which, by assumption, is possible only if $j \in I$). We conclude that

$$T := \{x_I : I \subseteq \{1,\ldots,n\}, |I| = m\}$$

is correct for $\pi_k(\mathbb{R}^m)$ with $k := n-m$, since $|T| = \binom{n}{m} = \dim \pi_k$ and

$$L_I : x \longmapsto \prod_{j\notin I} \frac{1 + a_j * x}{1 + a_j * x_I} \qquad \in \pi_k$$

with $L_I(x_J) = \delta_{IJ}$.

It is a nice exercise to develop a Newton form for the resulting polynomial interpolation scheme. This leads to a particular generalization of divided differences quite different from the tensor product construction with which we began this discussion.

The Newton approach has recently been generalized by Gasca & Maeztu [1980]. Although the idea is proposed in $\mathbb{R}^m$, I shall follow its authors and discuss details only in $\mathbb{R}^2$. Start with a straight line $1 + a_1 * x = 0$. (This is not quite the most general line, but that doesn't matter.) Add a bunch of lines $1 + a_{1i} * x = 0$, intersecting the first line at distinct points x_{1i}, $i=1,\ldots,m_1$. Form the polynomials

$$p_{1i}: x \longmapsto (1 + a_{11}{}^*x) \ldots (1 + a_{1,i-1}{}^*x) \ , \ i=1,\ldots,m_1 \ .$$

The $\mathrm{LIP}(\mathrm{span}(p_{1i}),\mathrm{span}([x_{1i}]))$ is correct since the Gram matrix $(p_{1j}(x_{1i}))$ is triangular with nonzero diagonal, hence invertible.

Now add a second line $1 + a_2{}^*x = 0$ intersecting the first line at a point other than the x_{1i}'s (if at all), and add a second bunch of lines $1 + a_{2i}{}^*x = 0$, intersecting the second line at distinct points x_{2i}, $i=1,\ldots,m_2$. Form the corresponding polynomials

$$p_{2i}: x \longmapsto (1 + a_1{}^*x)(1 + a_{2i}{}^*x) \ldots (1 + a_{2,i-1}{}^*x) \ , \ i=1,\ldots,m_2 \ .$$

Then (and this is the salient part of the construction), the single linear factor $(1 + a_2{}^*x)$ in the p_{2i} makes them vanish at all the earlier interpolation points x_{1j}. With this, the matrix $(p_{rj}(x_{si}))$ in lexicographic order is triangular with nonvanishing diagonal, and thus the $\mathrm{LIP}(\mathrm{span}(p_{rj}),\mathrm{span}([x_{si}]))$ is correct.

The general pattern is now clear. What is less clear is just what $\mathrm{span}(p_{rj})$ might be and, in more than two variables, things become horrendous. Still, for certain regular choices (see Maeztu [1982]), $\mathrm{span}(p_{rj})$ can be shown to coincide with π_k and the corresponding correct point set can be more general than the simplicial choice, but not more general than those covered by Theorem 4.1.

In his 1978 thesis (see Kergin [1978], [1980]), Paul Kergin proposes a totally different approach to multivariate polynomial interpolation which, in a way, gave impetus to all the material yet to be discussed in these lectures. I begin with Kergin's result as he stated it.

Theorem 4.2. *For any point sequence* $t_0, \ldots, t_n$ *in* $\mathbb{R}^m$ *there exists exactly one map* $P: C^{(n)}(\mathbb{R}^m) \longrightarrow \pi_n$ *so that*

(i) P *is linear;*

(ii) $\forall\, g \in C^{(n)}$ $\forall\, 0 \le k \le n$ $\forall\, q_k \in \pi_k$ *homogeneous of degree* k $\forall\, J \subseteq \{0,\ldots,n\}$ *with* $|J| = k+1$,

$$q_k(D)Pg = q_k(D)g$$

at some point in $\mathrm{conv}(t_j)_{j\in J}$.

Here, as earlier, conv T denotes the convex hull of the point set T. Further, $q_k \in \pi_k$ is homogeneous of degree k in case

$$q_k = \sum_{|\alpha|=k} A_\alpha (\,)^\alpha \ .$$

Consider for a moment the special case $m = 2$. The stated requirement for $k = 0$ forces Pg to agree with g at each of the t_i's. For $k = 1$, we have

$$q_k(D) = A_{(1,0)}D_1 g + A_{(0,1)}D_2 g \ .$$

Thus, if $t_i \neq t_j$, then Pg would have to match any such derivative $q_l(D)$ somewhere on the segment between t_i and t_j. This condition is already satisfied in case $q_l(D)$ is in the same direction as the segment since Pg matches g at t_i and t_j. Therefore, this imposes just one additional condition, viz. that the derivative normal to the segment be matched somewhere along the segment. If $t_i = t_j$ for $i \neq j$, we get <u>two</u> additional conditions, viz. that Pg have the same tangent plane at t_i as does g. In other words, we obtain osculatory interpolation.

An extreme case of osculatory interpolation occurs in case $t_0 = \ldots = t_n$. Now Pg is necessarily just the Taylor expansion of degree $\leq n$ for g at t_0.

Kergin begins the proof of his theorem with the observation that P is necessarily continuous on $X := C^{(n)}(G)$ for any bounded G containing $t_0, \ldots, t_n$. Since P is linear, this requires only to show that P is bounded (on X). This latter fact Kergin shows by observing that, by assumption, the leading coefficients of Pg agree with the corresponding normalized derivatives of g at certain points, hence can be bounded in terms of $\|g\|_X$. He then considers $Pg - Lg$, with Lg the leading terms of Pg just estimated, and observes that the leading terms of the resulting polynomial are of lower order and must interpolate to the corresponding normalized derivatives of $g - Lg$ at certain points, hence can be bounded in terms of $\|g - Lg\|_X$, therefore in terms of $\|g\|_X$. The inductive argument is now clear.

In consequence, P can be understood entirely from its action on a fundamental subset of $C^{(n)}$, i.e., a subset R whose finite linear combinations are dense in $C^{(n)}$. Kergin chooses R to consist of socalled plane waves (F. John) or ridge functions (C. A. Micchelli),

$$R := \{g \circ \lambda : g \in C^{(n)}(\mathbb{R}), \lambda \in (\mathbb{R}^m)'\} .$$

Such a function is constant in all planes normal to a certain direction. Explicitly,

$$(g \circ \lambda)(x) = g(\lambda^* x) = g\left(\Sigma_1^m \lambda(i)x(i)\right) .$$

Note that it is sufficient to take just one suitable g, e.g., $g: t \longmapsto e^{it}$.

Next, Kergin shows uniqueness. To be sure, the claim is not that, for a given g, there is a unique $Pg \in \pi_n$ satisfying the conditions described in (ii), for that is not true. For example, the function $g: x \longmapsto x(1)x(2)$ has all functions $p: x \longmapsto ax(2)$ with $0 \leq a \leq 1$ as

linear "interpolants" at the points $(0,0)$ and $(1,0)$ in that sense. Rather, Kergin claims the uniqueness of such a linear map P and proves it by showing that plane waves (with the univariate function g a polynomial) have unique "interpolants". Given the many conditions P has to satisfy, the uniqueness is not surprising. The hard part is to show existence.

For this, Kergin introduces (in rather different notation) the linear functionals

$$\int_{[x_0,\ldots,x_k]} g := \int_0^1 \ldots \int_0^{s_k} g(x_0 + s_1\nabla x_1 + \ldots + s_k \nabla x_k)\, ds_k \ldots ds_1 \qquad (4.1)$$

and sets

$$Q := \text{span}\{ g \longmapsto \int_{[t_J]} q_k(D)g \;:\; |J| = k+1,\ k=0,\ldots,n\}.$$

Here, $t_J := (t_j)_{j\in J}$, with $J \subseteq \{0,\ldots,n\}$ as before. Then, by some hard counting, Kergin shows that $\dim Q < \dim \pi_n$. Add to this the fact that

$$Q_\perp \cap \pi_n = \{0\} \qquad (4.2)$$

and you can conclude that the $\text{LIP}(\pi_n, Q)$ is correct. Now take for P the resulting projector. Then

$$\int_{[t_J]} q_k(D)(g - Pg) = 0,$$

hence $q_k(D)(g - Pg) = 0$ at some point in $\text{conv}\, t_J$. The claim (4.2) is established by an inductive argument: If $p \in \pi_n \cap Q_\perp$, then, for all $|\alpha| = n$, $\int_{[t_0,\ldots,t_n]} D^\alpha p = 0$, therefore $D^\alpha p = 0$, i.e., $p \in \pi_{n-1}$, etc.

Micchelli & Milman [1980] give a striking formulation of Kergin's interpolation scheme which shows it to be a "lifting" of the Newton form of the univariate interpolating polynomial. Micchelli came to this by noticing that Kergin's linear functionals (4.1) are closely related to the divided difference (as the notation used in (4.1) already intimates) via the **Hermite-Genocchi formula** (see Nörlund [1923; p.16]):

$$[\tau_0,\ldots,\tau_k]g = \int_{[\tau_0,\ldots,\tau_k]} D^k g$$

for any sufficiently smooth univariate g, a fact easily proved by induction. This allows us to write the univariate polynomial interpolant $P_\tau f$ in Newton form as

$$(P_\tau g)(x) = \Sigma_0^n (x-\tau_0)\ldots(x-\tau_{k-1}) \int_{[\tau_0,\ldots,\tau_k]} D^k g .$$

Also, recall that $[\tau_J]P_\tau g = [\tau_J]g$ for all $J \subseteq \{0,\ldots,n\}$. Now consider the Micchelli-Milman definition

$$Pf : x \longmapsto \sum_{k=0}^{n} [t_0,\ldots,t_k]\, D_{x-t_0} \cdots D_{x-t_{k-1}} f \tag{4.3}$$

for any $f \in X := C^{(n)}(\mathbb{R}^m)$. The resulting map P is linear and continuous on X, hence can be understood by looking at its action on the set R of plane waves. For $f = g\circ\lambda \in R$, one computes

$$D_y f = \Sigma_1^m y(i)\, g^{(1)}(\lambda^*\cdot)\lambda(i) = (\lambda^* y)\, g^{(1)}\circ\lambda ,$$

therefore

$$D_{x-t_0}\cdots D_{x-t_{k-1}} f = \lambda^*(x-t_0) \cdots \lambda^*(x-t_{k-1})\, g^{(k)}\circ\lambda$$

and so

$$Pf(x) = \Sigma_0^n \lambda^*(x-t_0) \cdots \lambda^*(x-t_{k-1}) \int_{[\lambda^* t_0,\ldots,\lambda^* t_k]} g^{(k)} .$$

The last integral equals $[\lambda^* t_0,\ldots,\lambda^* t_k]g$, by the Hermite-Genocchi formula. Therefore, finally,

$$P(g\circ\lambda) = (P_{(\lambda^* t_i)} g)\circ\lambda .$$

This is the crucial observation. It shows that $\operatorname{ran} P \subseteq \pi_n$ and that, for any $f = g\circ\lambda \in R$, any polynomial q_k homogeneous of degree k and any $J \subseteq \{0,\ldots,n\}$ with $|J| = k+1$,

$$\int_{[t_J]} q_k(D) Pf = \int_{[t_J]} q_k(D) f ,$$

since, for such an f,

$$q_k(D) f = \sum_{|\alpha|=k} A_\alpha D_1^{\alpha(1)} \cdots D_m^{\alpha(m)} f = \sum_{|\alpha|=k} A_\alpha \lambda^\alpha\, g^{(k)}\circ\lambda$$

and

$$\int_{[t_J]} (P_{(\lambda^* t_i)} g)^{(k)}\circ\lambda = [\lambda^* t_J] P_{(\lambda^* t_i)} g = [\lambda^* t_J] g .$$

This establishes that (4.3) is a formula for Kergin's map. |||

Micchelli [1980] offers additional detail, e.g., the error formula one associates with the Newton form which leads to a constructive proof of the Bramble-Hilbert lemma (see Bramble & Hilbert [1970]).

Kergin's scheme raises some questions. In contrast to its univariate antecedent, it requires derivative information even when all the interpolation points are distinct. This has been remedied recently in Hakopian [1981]$_1$ by the simple device of lowering the order of the derivatives appearing in the definition of the linear functionals which make up Kergin's interpolation conditions Q. One may also investigate other variants of Kergin's scheme. In particular, Kergin's scheme makes it very attractive to study the linear functionals (4.1), as a basis for a suitable definition of the divided differences of a function of several variables. This was, in fact, the consideration which led Kergin to his scheme in the first place.

The study of these linear functionals led C. A. Micchelli to the recurrence relations for multivariate B-splines and so opened up that fruitful area of research which is the topic of the remaining lectures.

5. Multivariate B-splines

Following the lead of Schoenberg [1965], the multivariate B-spline $M(\cdot|t_0,\ldots,t_n)$ was defined in de Boor [1976] by the rule

$$M(x|t_0,\ldots,t_n) := \frac{\mathrm{vol}_{n-m}(P^{-1}x)\cap[t_0,\ldots,t_n]}{\mathrm{vol}_n\,[t_0,\ldots,t_n]} \quad , \quad x \in \mathbb{R}^m \ ,$$

thus generalizing a particular characterization of the univariate B-spline due to Curry & Schoenberg [1966]. Here, $t_0, \ldots, t_n$ are points in $\mathbb{R}^n$, $[K]$ is the convex hull of the point set K, and P is the canonical projector

$$P : \mathbb{R}^n \longrightarrow \mathbb{R}^m : x \longmapsto (x(i))_1^m \ .$$

Further, $\mathrm{vol}_k(K)$ is the k-dimensional volume of the set K.

Such a B-spline is a nonnegative piecewise polynomial function of degree $\leq k := n - m$, its support is $[Pt_0,\ldots,Pt_n]$, and it is in C^{n-m-1} as long as the knots $Pt_0, \ldots, Pt_n$ are in general position. All this will be proved shortly.

In 1978, C. A. Micchelli [1980] proposed an equivalent but more suitable and flexible definition of $M(\cdot|t_0,\ldots,t_n)$ as the **distribution** on $C_0^\infty(\mathbb{R}^m)$ given by the rule

$$M(\cdot|t_0,\ldots,t_n) : f \longmapsto n! \int_{[t_0,\ldots,t_n]} f\circ P \ . \tag{5.1}$$

This definition makes sense even if the t_i's are not in general position. $M(\cdot|t_0,\ldots,t_n)$ is a function (in $L_\infty(\mathbb{R}^m)$) if and only if $\mathrm{vol}_m[Pt_0,\ldots,Pt_n] \neq 0$ and, in that case,

$$\int_{\mathbb{R}^m} M(\cdot|t_0,\ldots,t_n)f = n! \int_{[t_0,\ldots,t_n]} f\circ P \ . \tag{5.2}$$

More than that, Micchelli [1980] proved recurrence relations for these multivariate B-splines, of the following form.

Theorem 5.1. (i) If $x = \sum \alpha_i Pt_i$ with $\sum \alpha_i = 0$, then

$$D_x M(\cdot|t_0,\ldots,t_n) = n \sum \alpha_i M(\cdot|t_0,\ldots,t_{i-1},t_{i+1},\ldots,t_n) \ .$$

(ii) If $x = \sum \alpha_i Pt_i$ with $\sum \alpha_i = 1$, then

$$(n-m)\, M(x|t_0,\ldots,t_n) = n \sum \alpha_i M(x|t_0,\ldots,t_{i-1},t_{i+1},\ldots,t_n) \ .$$

These recurrence relations were proved almost simultaneously by a different approach by Dahmen [1979]$_1$ and have since then been given different proofs by Micchelli [1979], Höllig [1980], Hakopian [1980], de Boor & Höllig [1981] and perhaps others. I'll now give a version of this last proof, from de Boor & Höllig [1982], not only because, naturally, I like it best, but because it covers a more general situation than described so far.

To begin with, I have to clear up an inconsistency in the notation I have employed. This inconsistency shows up in (5.2) where both

$$\int_{\mathbb{R}^m} \quad \text{and} \quad \int_{[t_0,\dots,t_n]}$$

occur and $[t_0,\dots,t_n]$ is, off-hand, not meant as the convex hull of the points $t_0, \dots, t_n$, but rather as an indication that the integral is to be formed as described in (4.1). These two meanings only differ by a scalar factor, though, viz. the factor $n!\mathrm{vol}_n[t_0,\dots,t_n]$, hence could be made to coincide if, in (4.1), we multiplied the right hand side by $n!\mathrm{vol}_n\,[t_0,\dots,t_n]$. I settle this inconsistency instead by abandoning from now on the interpretation (4.1) and entirely rely on the more naive interpretation of

$$\int_{[t_0,\dots,t_n]}$$

as the integral over the convex hull of the points $t_0, \dots, t_n$.

Consider now, more generally, a polyhedral convex set B in $\mathbb{R}^n$, some linear map P into $\mathbb{R}^m$ and having B in its domain, and the distribution M_B on $\mathbb{R}^m$ defined by the rule

$$M_B : f \longmapsto \int_B f\circ P \quad , \tag{5.3}$$

i.e., as the P-shadow of B. The simplex spline $M(\cdot|t_0,\dots,t_n)$ results when $B = [t_0,..,t_n]/\mathrm{vol}_n[t_0,\dots,t_n]$ and P is the canonical map $\mathbb{R}^n \longrightarrow \mathbb{R}^m$, but it has already turned out to be fruitful to allow more general sets B and, usually as a simplification, more general linear maps P.

In any case, the recurrence relations can be proved in this generality. The relevant observation is that the boundary of such a convex polyhedral set consists again of convex polyhedral sets B_i, of one dimension lower, hence Stokes' Theorem can be used to relate $M := M_B$ to $M_i := M_{B_i}$. For this, we also need the corresponding outward normal ν_i to B at B_i relative to the affine hull of B, and an arbitrary point b_i in the affine hull of B_i. For simplicity, we assume that B is a body, i.e., B has $\mathbb{R}^n$ as its affine hull. With these assumptions and notations, the following theorem holds.

Theorem 5.2. (i) $D_{Pz}M = -\Sigma_i\, z^*\nu_i\, M_i$, all $z \in R^n$.
(ii) $(n-m)M(Pz) = \Sigma_i\, (b_i-z)^*\nu_i\, M_i(Pz)$, all $z \in R^n$.
(iii) $DM = (n-m)M - \Sigma_i\, b_i^*\nu_i\, M_i$.

The proof of (i) is immediate:

$$(D_{Pz}M)f = -\int_B (D_{Pz}f)\circ P = -\int_B D_z(f\circ P) = -\int_{\partial B} z^*\nu\, f\circ P$$
$$= -\Sigma_i \int_{B_i} z^*\nu_i\, f\circ P = -\Sigma_i\, z^*\nu_i\, M_i f .$$

This uses the fact that, by definition, the derivative D_yM of the distribution M is the distribution obtained by the rule $f \longmapsto M(-D_yf)$, and the standard interplay

$$D_y(f\circ P) = (D_{Py}f)\circ P$$

between differentiation and linear change of variables. This interplay also proves that

$$(Df)(Px) = (D_{Px}f)(Px) = (D_x(f\circ P))(x) = (D(f\circ P))(x) \qquad (5.4)$$

with D the first order differential operator given by the rule

$$Df := \sum_{j=1}^{r} F_j\, D_j f$$

in case f has its domain in R^r , and with

$$(F_jf)(x) := x(j)f(x) \text{ , all } j .$$

Thus $(Df)(x) = (D_xf)(x)$, and the adjoint of D is $-\Sigma\, D_jF_j$. This is of use in proving (iii): We have

$$D_jF_j = 1 + F_jD_j ,$$

therefore

$$-(DM)f = \int_B \Big(\sum_{j=1}^{m} D_jF_jf\Big)\circ P = mMf + \int_B (Df)\circ P$$

and

$$\int_B \sum_{i=1}^{n} D_iF_i(f\circ P) = nMf + \int_B D(f\circ P) .$$

Here, the last integral in the first line equals the last integral in the second, by (5.4). Therefore,

$$(DM)f = (n-m)Mf - \sum_{i=1}^{n} \int_B D_iF_i(f\circ P) .$$

This settles (iii) since

$$\sum_{i=1}^{n} \int_B D_iF_i(f\circ P) = \sum_{i=1}^{n} \int_{\partial B} \nu(i)\, F_i(f\circ P) = \int_{\partial B} (\cdot^*\nu)\, f\circ P$$

and, on the facet B_i , the function $(\cdot^*\nu)$ is constant.

Finally, to prove (ii), conclude from (i) and (iii) that, for any z with $Pz = x$,

$$0 = (D - D_{pz})M(x)$$
$$= (n-m)M(x) - \Sigma\, b_i{}^*\nu_i\, M_i(x) + \Sigma\, z^*\nu_i\, M_i(x) .$$

As an exercise, I specialize Theorem 5.2 to the situation of Theorem 5.1. This means that $B = [t_0,\dots,t_n]$ and that we may set $B_i := [(t_j)\setminus t_i]$, $i=0,\dots,n$. Then

$$M = \mathrm{vol}_n B\; M(\cdot|t_0,\dots,t_n) ,$$

therefore

$$M_i = \mathrm{vol}_{n-1}B_i\; M(\cdot|t_0,\dots,t_{i-1},t_{i+1},\dots,t_n) .$$

Also,

$$(t_i-b_i)^*\nu_i\ \mathrm{vol}\ B_i = -n\ \mathrm{vol}\ B$$

showing that the coeffients which appear in (i)-(iii) are, for this case, closely related to the barycentric or areal coordinates associated with the simplex B . In any case, if $z = \Sigma_j \alpha_j t_j$ with $\Sigma_j \alpha_j = 0$, then, since ν_i is perpendicular to the affine hull of $(t_j)\setminus t_i$, we have $(t_j-b_i)^*\nu_i = 0$ for all $j \neq i$ and so

$$z^*\nu_i = \Sigma_j \alpha_j (t_j-b_i)^*\nu_i = \alpha_i (t_i-b_i)^*\nu_i ,$$

therefore, from (i),

$$D_{pz}M(\cdot|(t_j)) = D_{pz}M/\mathrm{vol}\ B = (-\Sigma_i\, z^*\nu_i\, M_i)/\mathrm{vol}\ B$$
$$= n\ \Sigma_i\ \alpha_i\ M_i/\mathrm{vol}\ B_i = n\ \Sigma_i \alpha_i M(\cdot|(t_j)\setminus t_i)$$

This proves (i) of Micchelli's Theorem 5.1, with $x = Pz$. For Theorem 5.1(ii), we have, with $x = Pz$, $z = \Sigma_j \alpha_j t_j$ and $\Sigma_j \alpha_j = 1$, that

$$(b_i-z)^*\nu_i = (b_i - \Sigma \alpha_j t_j)^*\nu_i = \Sigma \alpha_j (b_i-t_j)^*\nu_i = \alpha_i (b_i-t_i)^*\nu_i ,$$

therefore, from (ii),

$$(n-m)M(x|(t_j)) = \Sigma_i (b_i-z)^*\nu_i\, M_i(x)/\mathrm{vol}\ B$$
$$= \Sigma_i\ \alpha_i\ (b_i-t_i)^*\nu_i\ M_i(x)/\mathrm{vol}\ B = \Sigma_i\ \alpha_i M(x|(t_j)\setminus t_i) .$$

W. Dahmen has pointed out to me that, once one recognizes that the recurrence relations in Theorem 5.1 can be written as in Theorem 5.2 in terms of facet normals, then Theorem 5.2 can be derived from Theorem 5.1 using the fact that any convex polyhedral body can be triangulated, i.e., is the essentially disjoint union of simplices, any two of which are either disjoint or else have exactly a face in common.

The recurrence relation (iii) was first stated and proved (for simplices) by Hakopian [1980], in his simple proof of Theorem 5.1.

Particularly useful choices for B other than the simplex include cones and boxes. The choice of the cone

$$\{t_0 + \Sigma_1^n \alpha_i(t_i - t_0) : 0 \leq \alpha_i \text{ , all } i\}$$

with vertex t_0 and generating rays $t_i - t_0$, $i=1,\ldots,n$ leads to the cone spline $M = M_B$. This is the truncated power introduced and heavily used in Dahmen [1979], in direct generalization of the functions

$$\mathbb{R} \longrightarrow \mathbb{R} : x \longmapsto (x - t)_+^{k-1}$$

familiar from univariate spline analysis. Of course, every such cone spline can be obtained as a translate of the P-shadow of the standard cone

$$\mathbb{R}_+^n := \{z \in \mathbb{R}^n : 0 \leq z(i) \text{ , all } i \}$$

for an appropriate choice of the linear map P .

Choice of the box or parallelepiped

$$\{t_0 + \Sigma_1^n \alpha_i(t_i - t_0) : 0 \leq \alpha_i \leq 1 \text{ , all } i\}$$

gives rise to the box spline introduced in de Boor & DeVore [1981] and further studied in de Boor & Höllig [1982]. Any such box spline is a translate of the P-shadow of the standard box

$$\{z \in \mathbb{R}^n : 0 \leq z(i) \leq 1 \text{ , all } i\}$$

for an appropriate choice of the linear map P .

Repeated application of the recurrence relation (i) provides the information that, for arbitrary vectors y_1 , ... , y_r ,

$$D_{y_1} \cdots D_{y_r} M_B \in \text{span}\{ M_F : F \text{ is an } (n-r)\text{dim.face of } B\}.$$

Thus all r-th order derivatives of M_B are in L_∞ provided PF is m-dimensional for every (n-r)-dimensional face F of B . This allows the following conclusions, asserted earlier for the simplex spline:

(iv) $M_B \in L_\infty^{(d)} \subseteq C^{(d-1)}$, <u>with</u>

$d := \max \{ r : \dim PF = m \text{ , } \forall \text{ (n-r)-dim.faces } F \text{ of } B\}$,

<u>and, for this</u> d , $M_B \notin C^{(d)}$.

(v) <u>If</u> PB <u>is m-dimensional, then</u> M_B <u>is a pp function of degree</u> $\leq k := n-m$, <u>with</u> $\operatorname{supp} M_B \subseteq PB$. Indeed, any (n-k-1)-dimensional face F of B is mapped by P into some hyperplane in $\mathbb{R}^m$, hence all (k+1)st order derivatives of M_B have their support entirely on some hyperplanes. This implies that, <u>on each connected component of the complement of</u>

$$\{ PF : F \text{ is a face of } B \text{ , } \dim PF < m \} \text{ ,}$$

M_B <u>agrees with some element of</u> π_k .

6. Approximation from the span of multivariate B-splines

A solitary B-spline is of little use in approximation. Thus we consider now a whole collection $B := (B)$ of polyhedral convex bodies in R^n, each giving rise to its P-shadow M_B, and ask just how B should be chosen so that we get a "useful" family (M_B). We use the well known properties of Schoenberg's univariate B-spline (see, e.g., de Boor [1976]) as a guide.

A first useful property of univariate B-splines is that, properly normalized, they form a partition of unity. This is not so hard to achieve with multivariate B-splines. We take for P the canonical map $R^n \to R^m$ and choose B as a partition of some slab $R^m \times \Delta$ in R^n, for some $\Delta \subseteq R^k$ (with $k := n-m$, as before). Then

$$\sum_B M_B(x) = \sum_B \mathrm{vol}_k P^{-1}x \cap B = \mathrm{vol}_k P^{-1}x \cap \bigcup_B B = \mathrm{vol}_k \Delta ,$$

i.e., we have a partition of the constant $\mathrm{vol}_k\Delta$. Choosing Δ to have $\mathrm{vol}_k\Delta = 1$ or else dividing each M_B by $\mathrm{vol}_k\Delta$, gives the desired partition of unity. We conclude at once that a continuous function f can be approximated from

$$S_B := \mathrm{span}\,(M_B)_{B\in B}$$

to within

$$\omega(f;|B|) ,$$

with

$$|B| := \sup_{B\in B} \mathrm{diam}\, PB .$$

The simple approximation

$$\sum f(\tau_B) M_B$$

with $\tau_B \in B$, all B, is that accurate (exercise).

Can we do better in case f is smoother? From the univariate theory, we would expect to get

$$\mathrm{dist}(f,S_B) = O(|B|^{k+1}) \qquad (6.1)$$

in case $f \in L_\infty^{k+1}$. This we could conclude at once if we had available a **quasi-interpolant** Q for S_B, i.e., a bounded linear map Q into S_B which is local and reproduces π_k. A typical specification of "local" would be to require that $Qf_{|C}$ depend only on $f_{|N(C)}$ with

$$N(C) := \bigcup\{\, PB : PB \cap C \neq \emptyset \,\} .$$

Then, as in Lebesgue's inequality, we could conclude that

$$f - Qf = f-p - Q(f-p) \text{ , for all } p \in \pi_k ,$$

therefore

$$\|(f-Qf)_{|C}\| \leq \|(f-p)_{|C}\| + \|Q\|\|(f-p)_{|N(C)}\| ,$$

and so

$$\|(f - Qf)|_C\| \leq (1 + \|Q\|) \operatorname{dist}_{N(C)}(f, \pi_k) .$$

This leads to (6.1).

A first requirement for such an argument is that $\pi_k \subseteq S_B$. This was established by Dahmen in $[1979]_2$ by a clever argument for the case that B is a triangulation of such a slab

$$S := \mathbb{R}^m \times \Delta ,$$

i.e., B consists of simplices. In effect, he deforms the slab appropriately to S' so that the function

$$p : x \mapsto \operatorname{vol}_k P^{-1}x \cap S'$$

is (locally) a polynomial of degree k , yet S' is still triangulated by the simplices B' , with B' the simplex into whose vertices the deformation sent the vertices of B . The deformation only takes place in directions perpendicular to $\mathbb{R}^m$. Therefore P carries the vertices of B and the corresponding vertices of B' to the same points. Consequently,

$$M_B/\operatorname{vol}_n B = M_{B'}/\operatorname{vol}_n B' ,$$

and so

$$p = \Sigma_{B \in \mathcal{B}} (\operatorname{vol} B'/\operatorname{vol} B)\, M_B .$$

He is able to modulate the deformation sufficiently to obtain a basis for π_k in this way.

This description neatly avoids discussion of some very nontrivial details.

In $[1980]_2$, Dahmen uses $\Delta = [0,1]^k$. There, he uses his result to support the claim that (6.1) holds (called Theorem 3.1 there), even without benefit of a quasi-interpolant, but I cannot follow the argument (even though it is a generalization of a univariate argument I once made up). I must therefore go the way of quasi-interpolants. Here I come across a difficulty first observed by Dahmen $[1980]_2$: With the choice $\Delta = [0,1]^k$, the B-splines (M_B) are linearly dependent. Dahmen overcomes this difficulty through an averaging process which provides a quasi-interpolant bounded in terms of a local mesh ratio.

There is another, practical difficulty, though, with this setup. The B-splines M_B , their support, their sets of discontinuities, all depend on the triangulation B of the slab S . In practice, one would like to start with some triangulation T of $\mathbb{R}^m$ and then construct B-splines in

$$\pi^r_{k,T} := \pi_{k,T} \cap C^{(r)} ,$$

the space of all pp functions of degree $\leq k$ and smoothness r

associated with that triangulation. Further, one would like to find enough B-splines to staff a basis for $\pi^r_{k,T}$, just as one is able to do for $m = 1$ (which acounts for the "B" in their name).

This, as it turns out, is too much to ask for, for various reasons. At present, the nature of the spaces $\pi^r_{k,T}$ is not at all understood, except when $r < 1$. We don't even know exactly when we are entitled to a local basis for that space. We must therefore be satisfied merely to construct $M_B \in \pi^r_{k,T}$. But even that is not possible in general, since the sets of discontinuities of M_B are associated with the projections of faces of B, hence not arbitrarily choosable. In particular, even if B is a simplex, the connected components of the complement of the discontinuity set need not be simplices.

One can hope, though, to obtain, for a given triangulation T, a linearly independent collection (M_B) of B-splines of prescribed smoothness, whose subdivision is a refinement of T and whose span S_B contains π_k. It is not difficult to do this for the univariate B-splines. Take for Δ the standard simplex $[e_0,\ldots,e_k]$ in $\mathbb{R}^k$ (with $e_0 := 0$, $e_j := (\delta_{ij})$ for $j > 0$) and triangulate $\mathbb{R}\times\Delta$ by the simplices

$$\tau_i := [t_i \times e_{[i]}, \ldots, t_{i+k} \times e_{[i+k]}], \quad \text{all } i,$$

with (t_i) the given knot sequence in $\mathbb{R}$ and $[j]$ the remainder on division of j by $k+1$. It is worthwhile to visualize this construction for $k = 1$ and 2. The construction uses the total ordering of $\mathbb{R}$ in an essential way, hence it is not easily generalized to the general case. Nevertheless, a construction of the desired type was found by Dahmen & Micchelli [1982] and independently by Höllig [1982]. The underlying idea is to get a B-spline basis (M_B) for $\pi_{k,T}$ and then to "pull apart their knots". Explicitly, with

$$\Delta = [e_0,\ldots,e_k]$$

the standard simplex in $\mathbb{R}^k$ and

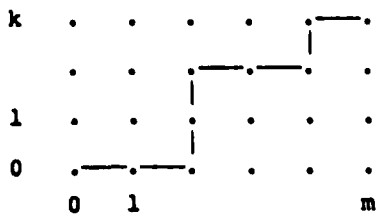
$$\tau := [v_0,\ldots,v_m]$$

any particular simplex in the triangulation T, the simploid

$$\tau\times\Delta$$

is triangulated using the combinatorial product $\tau\diamond\Delta$. I find it helpful to visualize this construction in the following way:

In the cartesian product $(v_i)\times(e_j)$, there are $\binom{m+k}{m}$ nondecreasing "paths" with endpoints (τ_0,e_0) and (τ_m,e_k). A typical "path" is shown in the following figure.



Associate with each such "path" the simplex σ in $\tau\times\Delta$ spanned by the points on that "path". It is a nice and worthwhile exercise to show that the resulting collection

$$\Sigma_\tau = (\sigma)$$

of simplices forms a triangulation for $\tau\times\Delta$. More than that, if we carry out this construction for each $\tau \in T$, making certain only that the ordering of the vertices of each τ used is consistent, i.e., comes from a total ordering of all the vertices appearing in T, then we obtain a triangulation

$$\Sigma := T\diamond\Delta$$

for $\mathbb{R}^m\times\Delta$. This is straightforward except, perhaps, for the assertion that

$$\sigma\cap\sigma' = [V_\sigma\cap V_{\sigma'}] \quad , \quad \text{all } \sigma, \sigma' \in \Sigma \ ,$$

with V_σ the set of vertices of the simplex σ.

Now consider $(M_\sigma)_{\sigma\in\Sigma}$. For each $\sigma \in \Sigma_\tau$, M_σ has support in τ, and $PV_\sigma = V_\tau$, hence M_σ has τ as its support and agrees with some polynomial there. For a given $\tau \in T$, there are exactly $\binom{m+k}{m}$ such M_σ, i.e., exactly enough to staff a basis for $\pi_{k|\tau}$. Since, by Dahmen's result, $\pi_k \subseteq S_\Sigma$, it follows that these M_σ form a basis for $\pi_{k|\tau}$. We conclude that $(M_\sigma)_{\sigma\in\Sigma}$ <u>is a basis for</u> $\pi_{k,T}$.

It's time to pull apart the knots. For this, denote by

$$V(\Sigma) = V(T) \times \{e_0,\ldots,e_k\}$$

the vertices of the triangulation $\Sigma = T\diamond\Delta$. Let

$$F : V(\Sigma) \longrightarrow \mathbb{R}^n : (v,e_i) \longmapsto (\overline{F}v,e_i) \ .$$

The only requirement Höllig [1982] makes on F (or $\overline{F}$) is that it be **locally finite**, i.e., $|\text{ran } F \cap C| < \infty$ for all bounded sets C. Dahmen & Micchelli [1982] require that $\text{sign det } \sigma = \text{sign det } F\sigma$, all $\sigma \in \Sigma$. Here,

$$\det \sigma$$

is the signed volume of the simplex σ. Its signature depends, of course, on the ordering of its vertices and the ordering is meant to be the ordering in $V(\Sigma)$.

With this, I am ready to state the basic result of this construction due to Dahmen & Micchelli and to Höllig.

Theorem 6.1. For all $y \in \mathbb{R}^m$,

$$(1 + x^*y)^k = \sum_{\sigma \in \Sigma} C_\sigma(y)\, M_{F\sigma}(x)$$

for all x at which all the simplex splines occuring on the right are continuous, with

$$C_\sigma(y) := k!\, \operatorname{sign}(\det \sigma)\, \det G_y F\sigma$$

and

$$G_y : \mathbb{R}^m \times \mathbb{R}^k \longrightarrow \mathbb{R}^n : (x,u) \longmapsto (x, (1+x^*y)u) \quad .$$

In effect, G_y carries out Dahmen's appropriate deformation of the slab $\mathbb{R}^m \times \Delta$ mentioned earlier.

Since y is arbitrary, we conclude that

$$\pi_k \subseteq S_{F\Sigma} \quad .$$

The quasi-interpolant is now immediate, in case F is not too violent. Specifically, assume that, for each $\tau \in T$, there is some ball b_τ which is contained in the support of every $M_{F\sigma}$ with $\sigma \in \Sigma_\tau$ and is outside the support of every other $M_{F\sigma}$. This implies that

$$S_{F\Sigma} = \operatorname{span} (M_{F\sigma})_{\sigma \in \Sigma_\tau} \quad \text{on} \quad b_\tau$$

and, as there are just enough simplex splines $M_{F\sigma}$ to staff a basis for π_k, this implies that

$$M_{F\sigma}|_{b_\tau} \in \pi_k|_{b_\tau} \quad \text{for every} \quad \sigma \in \Sigma_\tau \quad .$$

This allows the construction of linear functionals λ_σ as normpreserving extensions, to $L_1(b_\tau)$ say, of the coordinate functionals μ_σ on π_k which carry the polynomial $M_{F\sigma}|_{b_\tau}$ to $\delta_{\sigma\sigma'}$, all $\sigma, \sigma' \in \Sigma_\tau$. Since

$$\operatorname{supp} \lambda_\sigma \subseteq b_\tau \quad , \text{ all } \quad \sigma \in \Sigma_\tau \quad ,$$

this shows that (λ_σ) is dual to $(M_{F\sigma})$, i.e.,

$$\lambda_\sigma M_{F\sigma'} = \delta_{\sigma\sigma'} \quad , \text{ all } \quad \sigma \in \Sigma \quad .$$

The resulting quasi-interpolant

$$Q : f \longmapsto \sum \lambda_\sigma f\, M_{F\sigma}$$

is therefore even a linear projector onto $S_{F\Sigma}$, i.e., it reproduces all of $S_{F\Sigma}$ and not just π_k. The only concern is its size. $\|Q\|$ can be bounded in terms of the relative size of b_τ in τ. In particular,

all is well in case the vertex perturbation map F is not too violent.

Theorem 6.1 provides a generalization of **Marsden's identity**. It is fair to say that it is based on a **two-dimensional** version of Marsden's identity first proved by Goodman & Lee [1981]. These authors provide, in particular, the more explicit formula

$$C_\sigma(y) = k! \prod_{j=3}^{n} (1 + z_j * y) \quad ,$$

with the z_j determined as follows: The vertices v_j of $F\sigma$ are of the form

$$Pv_j \times e_i$$

and, for each $i = 0,\dots,k = n-2$, there is at least one j so that $v_j = (Pv_j, e_i)$. This leaves exactly two possibilities:

(i) for some i, there are three vertices of the form (Pv_j, e_i). Then we will call them v_0, v_1, v_2.

(ii) for two values of i, there are two vertices of the form (Pv_j, e_i). Then we will call one pair v_0, v_1 and the other v_2, v_3. With this, we take

$$z_j := \begin{cases} Pv_j & , j > 3 \\ \left\{ \begin{array}{ll} Pv_3 & , \text{if (i)} \\ \mathrm{aff}[Pv_0, Pv_1] \cap \mathrm{aff}[Pv_2, Pv_3] & , \text{if (ii)} \end{array} \right\} & , j = 3 \end{cases} .$$

Höllig [1982] gives a simple example to show that such a nice formula with linearly factored coefficients is, in general, not to be expected for $m > 2$. Still, for the practically important case $m = 2$, these simple formulae lead Goodman & Lee to the intriguing generalization

$$V : f \longmapsto \Sigma\, f(t_\sigma)\, M_{F\sigma}$$

of Schoenberg's variation diminishing spline operator. This operator V is obviously positive regardless of the choice of the t_σ. Goodman & Lee choose

$$t_\sigma := (z_3 + \dots + z_n)/(n-2)$$

since, in light of Theorem 6.1, this implies that V reproduces π_1. They are able to prove that, for any continuous f, Vf converges to f in the uniform norm as $|F\Sigma| \longrightarrow 0$, provided only that all the z_j's for σ lie in $PF\sigma$, an issue only for z_3 and only in case (ii).

There is an analogous quasi-interpolant construction for the span of certain translates of a **box spline** in de Boor & Höllig [1982]. The

arguments have a different flavor, though, since the resulting pp functions have regular meshes, hence are amenable to "cardinal spline" techniques familiar from Schoenberg [1973].

7. Epilogue

In these lectures, I have touched on only very few questions of current interest in multivariate approximation theory. Even if I restrict attention to splines and pp functions, there are several areas of current research which I had intended to discuss when I first prepared for these lectures but which, in the end, I did not manage to fit into the allotted time.

The nature of the space

$$S := \pi^r_{k,T}$$

of smooth pp functions on a given partition T of some $\Omega \subseteq \mathbb{R}^m$ is not at all understood. Questions of interest concern the existence of a locally supported basis for S, the dimension of S, the dimension of the subspace consisting of those $f \in S$ which vanish $(r+1)$-fold at the boundary of Ω, the degree of approximation achievable from S. There is the conjecture that sufficiently smooth functions can be approximated to within $O(|T|^s)$ if and only if S contains a local partition of every $p \in \pi_{s-1}$, but attempts to prove this by construction of a quasi-interpolant have required, in addition, some kind of stability of the partitions. Work concerning $\dim S$ has been done only for $m = 2$ and initially only for $r = 1$, the first nontrivial case. See Strang [1974], Morgan & Scott [1975], and the survey of Schumaker [1979]. Most recently, Chui & Wang [1981]$_{1-3}$ have given precise results for certain T and arbitrary r. The existence of a local partition of unity in S is taken up in de Boor & DeVore [1981] for certain regular T in order to understand better the degree of approximation from S. These questions are further pursued in de Boor & Höllig [1982]. In both papers, the relationship between S and B-splines in S is explored, but even in the context of a simple and regular T (e.g., a rectangular grid with all north-east diagonals drawn in), this relationship is not yet fully understood.

The adaptive choice of the partition T is the topic of de Boor & Rice [1979]. Dahmen [1982] describes one way to use simplex splines adaptively. The degree of approximation achievable from S by proper

choice of T is the topic of Dahmen, DeVore & Scherer [1980].

Practial aspects of approximation from $S = \pi^r_{k,T}$ on $\mathbf{R}^2$ are the topic of the two survey papers Barnhill [1977] and Schumaker [1976]. An interesting comparison of methods is given in Franke [1982]. And then there is the vast literature on the constructive aspects of the **Finite Element Method**! Some references of particular interest to Approximation Theory are : Ciarlet & Raviart $[1972]_{1,2}$, Courant [1943], Fix & Strang [1969], Guglielmo [1969] , Strang & Fix [1973], but this is clearly just a taste.

The variational approach to splines is, of course, not restricted to the univariate situation. Already Golomb & Weinberger [1959] consider particular bivariate examples as illustrations of the general theory. This theory has the following setting. A collection Λ of continuous linear functionals on some linear space X is given. Since the problem will only involve

$$\Lambda_\perp := \{x \in X : \lambda x = 0 \text{ for all } \lambda \in \Lambda\} ,$$

we might as well assume that Λ is a closed subspace of X^*. Further, a bounded linear map T from X to some normed linear space Y is given. The problem is to determine, for given $x \in X$, if possible, an element x^* at which the map

$$x + \Lambda_\perp \longrightarrow \mathbf{R}_+ : y \longmapsto \|Ty\|$$

takes on its minimum. Such a minimizer x^* is called a (T,Λ)-**spline interpolant** to x . The word "interpolant" is appropriate since x^* agrees with x on Λ . Golomb & Weinberger [1959] deal with the special case: $X = Y$ is a Hilbert space and $T = 1$. In this setting, the map $x \longmapsto x^*$ is just the orthogonal projector onto Λ considered as a subspace of X . This schizophrenic nature of the interpolation conditions Λ , linear functionals on the one hand and elements of X on the other, is at the heart of the practical application of this theory. In standard Hilbert spaces of smooth functions on some domain Ω , the linear functional $[t]$ of evaluation at some point t turns into the section $G(t,\cdot)$ of the appropriate Green's function. In particular, when $X = L_2^{(k)}[a,b]$, then $[t]$ is represented by a pp function of degree $2k-1$ and in $C^{(2k-2)}$ with just one breakpoint, at t .

I was held back from exploring multivariate splines obtained in this way by the realization that this would require me to obtain and work with the Green's function relevant to X . This would usually not be polynomial nor locally simple, and would depend essentially on the

domain Ω . Duchon [1976], [1977] dealt with such objections by the very effective device of choosing all of R^m for Ω . The resulting thin plate splines have already found practical use. Their theory is described invitingly in Meinguet [1979].

References

N. I. Akhiezer [1967], Vorlesungen über Approximationstheorie, 2., verbesserte Auflage, Akademie-Verlag, Berlin; appeared also as Theory of Approximation, F. Ungar Publ., New York, 1956.

R. E. Barnhill [1977], Representation and approximation of surfaces, in Mathematical Software III, J. Rice ed., Academic Press, New York, 69-120.

C. de Boor [1976], Splines as linear combinations of B-splines, in Approximation Theory II, G.G. Lorentz, C.K. Chui & L.L. Schumaker eds., Academic Press, 1-47.

C. de Boor [1979], Efficient computer manipulation of tensor products, ACM Trans.Math.Software 5, 173-182.

C. de Boor & R. DeVore [1981], Approximation by smooth multivariate splines, Math.Research Center TSR #2319. Trans.Amer.Math.Soc., to appear.

C. de Boor & K. Höllig [1981], Recurrence relations for multivariate B-splines, Math.Research Center TSR #2215. Proc.Amer.Math.Soc., to appear.

C. de Boor & K. Höllig [1982], B-splines from parallelepipeds, Math.Research Center TSR #2320.

C. de Boor & J. R. Rice [1979], An adaptive algorithm for multivariate approximation giving optimal convergence rates, J.Approx. Theory 25, 337-359.

J. H. Bramble & S. R. Hilbert [1970], Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation, SIAM J.Numer.Anal. 7, 112-124.

J. H. Bramble & S. R. Hilbert [1971], Bounds for a class of linear functionals with applications to Hermite interpolation, Numer.Math. 16, 362-369.

E. W. Cheney [1966], Introduction to Approximation Theory, McGraw-Hill, New York.

C. K. Chui & R.-H. Wang [1981]$_1$, Multivariate B-splines on triangulated rectangles, CAT # 6, Center for Approximation Theory, Texas A&M University, College Station, TX.

C. K. Chui & R.-H. Wang [1981]$_2$, On a bivariate B-spline basis, CAT # 7.

C. K. Chui & R.-H. Wang [1981]$_3$, Multivariate spline spaces, CAT #9.

K. C. Chung & T. H. Yao [1977], On lattices admitting unique Lagrange interpolations, SIAM J.Numer.Anal. 14, 735-741.

P. G. Ciarlet & R. A. Raviart [1972]$_1$, General Lagrange and Hermite interpolation in R^k with applications to finite element methods, Arch.Rat.Mech.Anal. 46,177-199.

P. G. Ciarlet & R. A. Raviart [1972]$_2$, Interpolation theory over curved elements, with applications to finite element methods, Computer Methods in Appl.Mech.Eng. 1, 217-249.

R. Courant [1943], Variational methods for the solution of problems of equilibrium and vibrations, Bull.Amer.Math.Soc. 49, 1-23.

H. B. Curry & I. J. Schoenberg [1966], Pólya frequency functions IV. The fundamental spline functions and their limits, J.d'Anal.Math. 17, 71-107.

W. Dahmen [1979]$_1$, Multivariate B-splines - recurrence relations and linear combinations of truncated powers, in Multivariate Approximation Theory, W. Schempp & K. Zeller eds., Birkhäuser, Basel, 64-82.

W. Dahmen [1979]$_2$, Polynomials as linear combinations of multivariate B-splines, Math.Z. 169, 93-98.

W. Dahmen [1980]$_1$, On multivariate B-splines, SIAM J.Numer.Anal. 17, 179-191.

W. Dahmen [1980]$_2$, Approximation by linear combinations of multivariate B-splines, J.Approx.Theory, to appear.

W. Dahmen [1982], Adaptive approximation by multivariate smooth splines, J.Approx.Theory, to appear.

W. Dahmen, R. DeVore & K. Scherer [1980], Multi-dimensional spline approximation, SIAM J.Numer.Anal. 17, 380-402.

W. Dahmen & C. A. Micchelli [1980], On limits of multivariate B-splines, Math.Research Center TSR # 2114. J.d'Anal.Math. 39 (1981),256-278.

W. Dahmen & C. A. Micchelli [1982], On the linear independence of multivariate B-splines.I, Triangulation of simploids, SIAM J.Numer.Anal. xx, xxx-xxx.

P. Davis [1963], Interpolation and Approximation, Blaisdell, Waltham MA. Now available from Dover.

J. Duchon [1976], Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, R.A.I.R.O. Analyse numerique 10, 5-12.

J. Duchon [1977], Splines minimizing rotation-invariant seminorms in Sobolev spaces, in Constructive Theory of Functions in Several Variables, Oberwolfach 1976, W. Schempp & K. Zeller eds., Springer-Verlag, Heidelberg, 85-100.

G. Fix & G. Strang [1969], Fourier analysis of the finite element method in Ritz-Galerkin theory, Studies in Appl.Math. 48, 265-273.

R. Franke [1982], Testing methods for scattered data interpolation and some results, Math.Comp. 38, 181-200.

M. Gasca & J. I. Maeztu [1980], On Lagrange and Hermite interpolation in R^K, Numer.Math., to appear.

F. di Guglielmo [1969], Construction d'approximations des espaces de Sobolev sur des reseaux en simplexes, Calcolo 6, 279-331.

M. Golomb & H. F. Weinberger [1959], Optimal approximation and error bounds, in On Numerical Approximation, R. Langer ed., U. of Wisconsin Press, 117-190.

T. N. T. Goodman & S. L. Lee [1981], Spline approximation operators of Bernstein-Schoenberg type in one and two variables, J.Approx.Theory 33, 248-263.

W. J. Gordon [1969]$_1$, Spline-blended surface interpolation through curve networks, J.Math.Mech. 18, 931-952.

W. J. Gordon [1969]$_2$, Distributive lattices and approximation of multivariate functions, in Approximations with special emphasis on spline functions, I. J. Schoenberg ed., Academic Press, New York, 223-277.

H. Hakopian [1980], On multivariate B-splines, SIAM J.Numer.Anal. xx, xxx-xxx.

H. Hakopian [1981]$_1$, Les differences divisees de plusieurs variables et les interpolations multidimensionelles de types lagrangien et hermitien, C.R. Acad.Sc.Paris 292 , 453-456.

H. Hakopian [1981]$_2$, Multivariate spline functions, B-spline basis and polynomial interpolations, ms.

K. Höllig [1981], A remark on multivariate B-splines, J.Approx. Theory 33, 119-125.

K. Höllig [1982], Multivariate splines, SIAM J.Numer.Anal. xx, xxx-xxx.

P. Kergin [1978], Interpolation of C^k functions, Ph.D. Thesis, University of Toronto, Canada.

P. Kergin [1980], A natural interpolation of C^k functions, J.Approx.Theory 29, 278-293.

A. N. Kolmogorov [1936], Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse, Ann.Math. 37, 107-111.

G. G. Lorentz [1966], Approximation of Functions, Holt, Rinehart & Winston, New York.

J. I. Maeztu [1982], Divided differences associated with reversible systems in R^2, SIAM J.Numer.Anal., to appear.

J. Meinguet [1979], Multivariate interpolation at arbitrary points made simple, J.Appl.Math.Phys. (ZAMP) 30, 292-304.

C. A. Micchelli [1980], A constructive approach to Kergin interpolation in R^k: multivariate B-splines and Lagrange interpolation, Rocky Mountains J.Math.10, 485-497.

C. A. Micchelli [1979], On a numerically efficient method for computing multivariate B-splines, in Multivariate Approximation Theory, W. Schempp & K. Zeller eds., Birkhäuser, Basel, 211-248.

C. A. Micchelli & P. Milman [1980], A formula for Kergin interpolation in R^k, J.Approx.Theory 29, 294-296.

J. Morgan & R. Scott [1975], The dimension of the space of C^1 piecewise polynomials, ms.

P. D. Morris & E. W. Cheney [1974], On the existence and characterization of minimal projectors, J.reine angew.Math. 270, 61-76.

N. E. Nörlund [1924], Vorlesungen über Differenzenrechnung, Springer Grundlehren Bd. 13, Berlin.

M. J. D. Powell [1981], Approximation theory and methods, Cambridge University Press.

J. R. Rice [1964, 1969], The Approximation of Functions. Vols. I., II. , Addison-Wesley, Reading MA.

T. Rivlin [1969], In Introduction to the Approximation of Functions, Blaisdell, Waltham MA.

I. J. Schoenberg [1965], letter to Philip J. Davis dated May 31, 1965.

I. J. Schoenberg [1973], Cardinal Spline Interpolation, SIAM, Philadelphia PA.

A. Schönhage [1971], Approximationstheorie, de Gruyter, Berlin.

L. L. Schumaker [1976], Fitting surfaces to scattered data, in Approximation Theory II, G.G. Lorentz, C. K. Chui & L. L. Schumaker eds., Academic Press, New York, 203-268.

L. L. Schumaker [1979], On the dimension of spaces of piecewise polynomials in two variables, in Multivariate Approximation Theory, W. Schempp & K. Zeller eds., Birkhäuser, Basel, 396-412.

G. Strang [1974], The dimension of piecewise polynomials, and one-sided approximation, Springer Verlag Lecture Notes 363, 144-152.

G. Strang & G. J. Fix [1973], An Analysis of the Finite Element Method, Prentice-Hall, Englewood Cliffs.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2379	2. GOVT ACCESSION NO. AD-A116 248	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Topics in Multivariate Approximation Theory		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) C. de Boor		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 3 - Numerical Analysis and Computer Science
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE May 1982
		13. NUMBER OF PAGES 40
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

multivariate, B-splines, Kergin interpolation, linear projectors.

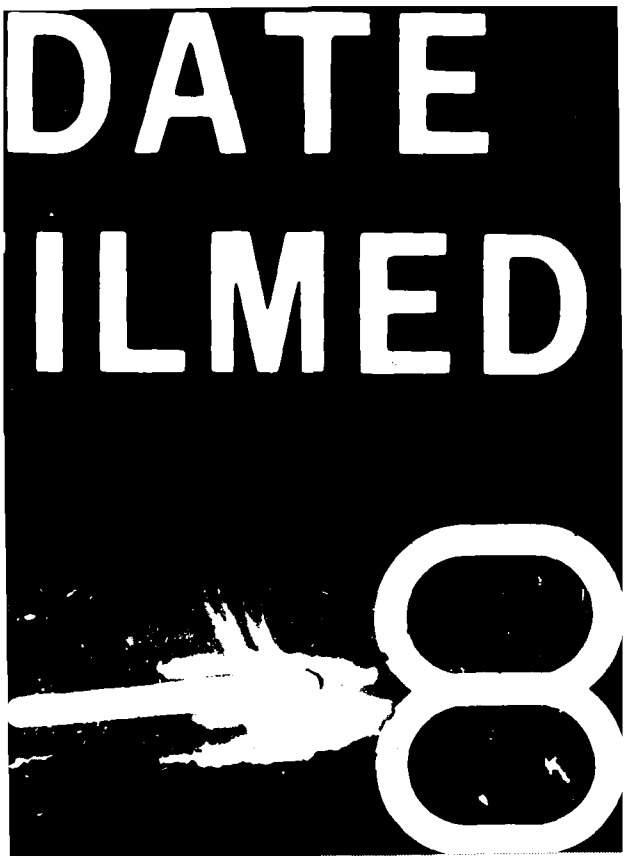
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Lectures delivered at the S.R.C. Numerical Analysis Summer School and Workshop at the University of Lancaster, England, July 19 - August 20, 1981. Topics include tensor products, multivariate polynomial interpolation, esp. Kergin Interpolation, and the recent developments of multivariate B-splines.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


DATE
ILMED
8